P1 1160885

THE UNITED STATES OF AMERICA

TO ALL TO WHOM THESE PRESENTS SHALL COME:

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office

April 26, 2004

THIS IS TO CERTIFY THAT ANNEXED HERETO IS A TRUE COPY FROM THE RECORDS OF THE UNITED STATES PATENT AND TRADEMARK OFFICE OF THOSE PAPERS OF THE BELOW IDENTIFIED PATENT APPLICATION THAT MET THE REQUIREMENTS TO BE GRANTED A FILING DATE.

APPLICATION NUMBER: *60/445,539*
FILING DATE: *February 06, 2003*
RELATED PCT APPLICATION NUMBER: *PCT/US04/03545*

By Authority of the
COMMISSIONER OF PATENTS AND TRADEMARKS

M. Sias

M. SIAS
Certifying Officer

PRIORITY DOCUMENT
SUBMITTED OR TRANSMITTED IN COMPLIANCE WITH RULE 17.1(a) OR (b)

02-10-03

60445539

PROVISIONAL APPLICATION COVER SHEET

This is a request for filing a PROVISIONAL APPLICATION under 37 CFR 1.53(b)(2).

Docket Number	14-03P	Type a plus sign (+) inside this box -	+

INVENTOR(s) APPLICANT(s)

LAST NAME	FIRST NAME	MIDDLE INITIAL	RESIDENCE (CITY AND EITHER STATE OR FOREIGN COUNTRY)
Mohnen	Debra		Athens, GA
Sterling	Jason	D	Athens, GA
Doong	Ron	L	Athens, GA
Kolli	Venkata	S.K.	Duluth, GA

TITLE OF THE INVENTION (280 characters max)

DISCOVERY OF A GENE ENCODING AN EXPRESSED ALPHA-1,4-GALACTURONOSYLTRANSFERASE IN *ARABIDOPSIS THALIANA* AND OF A FAMILY OF PROPOSED GALACTURONOSYLTRANSFERASES, AND THEIR GENES, IN *ARABIDOPSIS THALIANA* AND PLANTS

CORRESPONDENCE ADDRESS

23713
PATENT TRADEMARK OFFICE

GREENLEE, WINNER AND SULLIVAN, P.C.
5370 Manhattan Circle, Suite 201
Boulder, CO 80303 USA
Phone: (303) 499-8080; Fax (303) 499-8089
Email: winner@Greenwin.com

ENCLOSED APPLICATION PARTS (*check all that apply*)

X Specification Parts I & II 60 Number of Pages X Applicants Claim Small Entity Status. See 37 CFR 1.27
___ Drawing(s) ___ Number of Sheets X Other (specify) Appendices I & II 41 pages

METHOD OF PAYMENT (*check one*)

X A check or money order is enclosed to cover the Provisional filing fees.
___ The Commissioner is hereby authorized to charge filing fee and credit Deposit Account Number: 07-1969
X The Commissioner is hereby authorized to charge any deficiency or credit any overpayment to Deposit Account 07-1969

PROVISIONAL FILING FEE AMOUNT $80.00

The invention was made by an agency of the United States Government or under a contract with an agency of the United States Government.

___ No X Yes, the name of the U.S. Government agency and the Government contract number are:
USDA 98-35304-6772 and USDA 2001-03351

Respectfully submitted,

SIGNATURE Heeja Yoo-Warren

REGISTRATION NO. 45,495

DATE: February 6, 2003

CERTIFICATE OF MAILING
I hereby certify that this correspondence is being deposited with the United States Postal Service with sufficient postage as EXPRESS MAIL in an envelope addressed to the Assistant Commissioner for Patents, Washington, D.C. 20231.

6 February 2003 B. Kroge
Date B. Kroge
Express Mail Receipt No. EL 946 759 260 US

PROVISIONAL PATENT APPLICATION

Inventors:

Debra Mohnen
Jason Dwight Sterling
Ron Lou Doong
Venkata Siva Kumar Kolli

DISCOVERY OF A GENE ENCODING AN EXPRESSED ALPHA-1,4-GALACTURONOSYLTRANSFERASE IN *ARABIDOPSIS THALIANA* AND OF A FAMILY OF PROPOSED GALACTURONOSYLTRANSFERASES, AND THEIR GENES, IN *ARABIDOPSIS THALIANA* AND PLANTS

Prepared by:

GREENLEE, WINNER AND SULLIVAN, P.C.
5370 Manhattan Circle
Suite 201
Boulder, Colorado 80303
(303) 499-8080
FAX: (303) 499-8089

Attorney Docket No. 14-03P

14-03P

PART I

Role of Galacturonosyltransferase I (GALAT1) in Pectin Synthesis

Summary

Pectin is a complex family of cell wall polysaccharides that have multiple roles in plant growth, development and disease resistance. Pectins are also gelling and stabilizing agents in the food industry, and neutraceuticals and potential pharmaceuticals with beneficial effects on human health. We recently identified the first gene (*GALAT1*) for a pectin biosynthetic glycosyltransferase: *Arabidopsis thaliana* α-1,4-galacturonosyltransferase I (GALAT1). The specific aims of this proposal are to: **(1) Characterize heterologously expressed GALAT1 for kinetic constants, substrate and acceptor specificity, pH optima, thermal stability, quaternary structure, and post-translation glycosylation.** These characteristics will aid in determining the role of GALAT1 in homogalacturonan synthesis, possible role in rhamnogalacturonan I and/or rhamnogalacturonan II synthesis, and provide enzyme kinetic and specificity data for use in deciphering enzyme and regulatory mechanisms. **(2) Produce a series of mutated GALAT1 proteins by heterologous expression of site-directed mutated *GALAT1* genes and characterize the heterologously expressed mutated proteins.** Characterization of the kinetic constants of the mutated enzymes will facilitate the identification of amino acids important for enzyme activity and substrate/acceptor binding. **(3) Generate and use anti-GALAT1 antibodies for immunocytochemistry analysis of the subcellular location of GALAT1.** These experiments will establish in which region of the Golgi GALAT1 is located and allow the development of models for the compartmentalization of pectin synthesis within the endomembrane system. **(4) Use RNAi to generate GALAT1 knockouts and characterize the phenotype of the mutated plants.** The long-term goal is to understand the role of GALAT1 in pectin synthesis and use this knowledge to produce plants with improved agronomical value.

SPECIFIC AIMS

The goals of this proposal are to biochemically characterize *GALAT1*: the first pectin biosynthetic glycosyltransferase, and first galacturonosyltransferase from any species, to be at identified at the gene level. Our preliminary data suggest that the protein encoded by *GALAT1* (GALAT1), is the same α1,4-galacturonosyltransferase that has been shown by us[1-9] and others [10-15] to transfer galacturonic acid (GalA) from UDP-GalA onto homogalacturonan (HGA) (Fig. 1) and/or commercially available pectin acceptors. Thus, we propose that *GALAT1* encodes UDP-galacturonic acid:HGA α1,4-galacturonosyltransferase (i.e. GALAT1 in Table I). However, whether GALAT1 adds GalA only onto HGA, or can also add GalA to the backbone of the other pectic polysaccharides rhamnogalacturonan I (RG-I) (Fig. 2) or rhamnogalacturonan II (RG-II) (Fig. 3) is not known. Answers to these questions are important to identify which of the 4-9 different GalATs required for pectin synthesis (Table I) is encoded by *GALAT1*. Furthermore, since *GALAT1* is the first galacturonosyltransferase gene to be identified from any species, nothing is known about the regions of the protein involved in substrate or acceptor binding, nor are the catalytic amino acids known. The above questions will be addressed by Specific Aims 1

and 2 which are to characterize the kinetics and substrate specificity of native GALAT1 (Aim 1), and modified GALAT1 generated by site-directed mutagenesis (Aim 2) in order to define the regions of the enzyme important in catalysis and substrate binding and to determine the specificity of GALAT for HGA, RG-I and RG-II synthesis. The enzymes will also be analyzed for pH optima and thermal stability, to compare GALAT1 with the previously published GalATs (Table II), and possible quaternary structure and post-translation glycosylation. The results of the kinetic analyses will indicate whether the apparent biphasic kinetics obtained for GalAT in membrane bound and solubilized plant fractions (see below) are a characteristic of GALAT1 and, if so, whether the kinetics represent allosteric regulation of the enzyme. Specific Aim 3 is to determine the subcellular location of GALAT1 via immunocytochemistry using anti-GALAT1 antibodies to be generated. We previously showed that GalAT activity is membrane bound and localized to the lumen of the Golgi, and thus, proposed that the HGA biosynthetic GalAT is a Golgi resident enzyme with its catalytic site in the Golgi lumen. The predicted topology of GALAT1 encoded by *GALAT1* is completely consistent with these results since it is predicted to be a Type II membrane protein[16] with a an N-terminal cytosol facing region, a single transmembrane-spanning region and a larger C-terminal region. Location in the Golgi is also consistent with prior autoradiographic[17-20], immunocytochemical[21-23] and subcellular fractionation[7,24-26] results showing that pectin synthesis occurs in the Golgi. However, the distribution of the different pectin biosynthetic enzymes within the Golgi complex is not known. Such information is critical to determine which of the expected 72 predicted pectin biosynthetic enzymes[9] (see also Appendix 1) interact together to synthesize pectin. The proposed immunocytochemistry study would be the first information on the subcellular location of any pectin biosynthetic glycosyltransferase and would allow formulation of models for how pectin is synthesized within the Golgi. The antibodies to be generated in Specific Aim 3 would also be used to determine if GalAT is a multimeric protein, and if so, to establish its stoichiometry. Specific Aim 4 is to generate GALAT1 knockouts through RNAi to gain information about the role of GALAT1 *in vivo*. The results of these experiments will be most meaningful if viable plants that express no GALAT1 are obtained, and if such plants show changes in pectin structure. The resulting changes should represent the accumulation of GALAT1 precursors and an absence of the pectin normally generated by GALAT1. If time allows, additional experiments will also be initiated to use the anti-GALAT1 antibodies generated in Specific Aim 3 as "binding partners" to identify *Arabidopsis* proteins that bind to GALAT1 using immunoabsorption, followed by mass spectrometry sequencing of the proteins. This approach would allow us to begin to identify members of a predicted GALAT1 enzyme complex. A second line of research that would be initiated if time and resources allow is to attempt to obtain 3D structural information on GALAT1 via X-ray crystallography and possibly NRM spectroscopy. Our long term goal is to identify the enzymes involved in pectin synthesis, concentrating initially on HGA synthesis, and to decipher the process by which these enzymes work together to synthesize the complicated family of cell wall polysaccharides known as pectin.

Table I. *List of galacturonosyltransferase activities predicted to be required for pectin biosynthesis*[9]

Type of GalAT	Working[1] Number	Parent polymer[2]	Enzyme[3] Acceptor substrate *Enzyme activity*	Ref for Structure
D-GalAT	1	HGA	*GalAα1→4GalA *α1,4-GalAT*	[27]
D-GalAT	2	RG-I	L-Rhaα1→4GalA *α1,2-GalAT*	[27-29]

D-GalAT	3	RG-II	L-Rhaβ1→3Api*f* *α1,2-GalAT*	30,31
D-GalAT	4	RG-II	L-Rhaβ1→3Api*f* *β1,3GalAT*	30,31
D-GalAT	5 ?[4]	RG-I/HGA	GalAα1→2LRha *α1,4-GalAT*	---
D-GalAT	6 ?	RG-II/HGA	GalAα1→4GalA *α1,4-GalAT*	---
D-GalAT	7?	XGA	GalAα1→4(Xyl β1→3)GalA[5] *α1,4-GalAT*	32-35
D-GalAT	8 ?	AGA	GalAα1→4(Api*f* β1→2)GalA *α1,4-GalAT*	36,37
D-GalAT	9 ?	AGA	GalAα1→4(Api*f* β1→3)GalA *α1,4-GalAT*	36,37

[1]Numbers for different members of the same groups are given based on pectin structure and on the assumption that HGA is synthesized first, followed by RG-I and RG-II. The numbers were given[9] to facilitate a comparison of the enzymes, but final numbering will likely correspond to the order in which the genes are identified.

[2]HGA: homogalacturonan; RG-I: Rhamnogalacturonan I; RG-II: Rhamnogalacturonan II; XGA: Xylogalacturonan; AGA; Apiogalacturonan

[3]All sugars are D sugars and have pyranose rings unless otherwise indicated. Glycosyltranferases add to the glycosyl residue on the left* of the indicated acceptor.

[4]The ? means the designated GalAT may be required if a different GalAT in the list does not perform the designated function

[5]Glycosylresidue in the parenthesis is branched off the first GalA

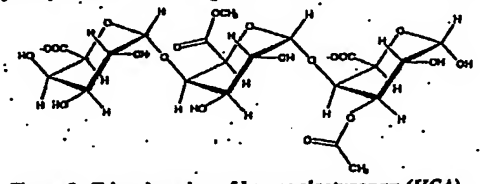


Figure 1. *Trimeric region of homogalacturonan (HGA).* HGA is a linear homopolymer of α-1,4-linked galacturonic acid that may be methylesterfied at C6 and acetylated at O2 or O3. Substituted galacturonans, such as RG-II, apiogalacturonan and xylogalacturonan have an HGA backbone.

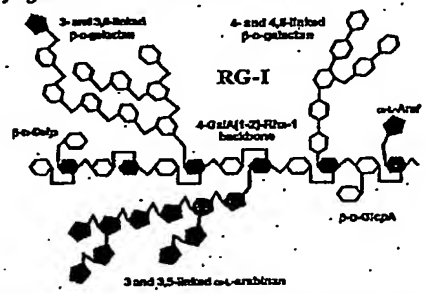


Figure 2. *Representative structure of rhamnogalacturonan I (RG-I).* RG-I has an alternating [→4)-α-D-GalpA-(1→2)- α-L-Rhap-(1→] backbone in which roughly 20-80% of the rhamnoses are substituted by arabinans, galactans, or arabinogalactans.

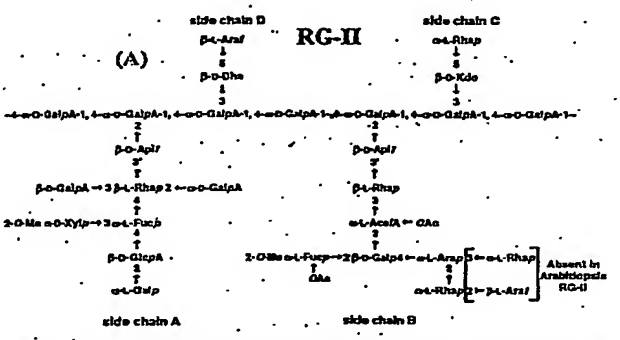


Figure 3. *Representative structure for rhamnogalacturonan II* (RG-II). RG-II has a backbone of 1,4-linked α-D-GalpA residues. GalA residues are also present in RG-II side chain B.

BACKGROUND

Pectin Structure & Function

Pectin is the most complex polysaccharide in the plant cell wall. It is comprises 30-40% of the primary wall of dicots and non-graminaceous monocots, and ~ 10% of the primary wall in the grass family. Pectins are a family of polysaccharides[6,8,27] that include homogalacturonan (HGA) (Fig. 1), rhamnogalacturonan-I (RG-I) (Fig. 2) and rhamnogalacturonan II (RG-II) (Fig. 3) as well xylogalacturonans (XGA)[32,34,38] and apiogalacturonans3[6,37] (see Appendix I). While the specific structure of each of these polysaccharides differs (Figs. 1-3), they are grouped into one family since they appear to be linked to each other in the wall and they each contain α-D-galacturonic acid connected in a 1,4-linkage.

HGA is the most abundant pectic polysaccharide, accounting for ~55%-70% of pectin[32]. HGA is a linear homopolymer of α1,4-linked D-galactosyluronic acid that is partially methylesterified at the C6 carboxyl group and may be partially acetylated at O-2 and/or O-3[8] (Fig. 1). Some plants also contain HGA that is substituted at the 2 or 3 position by D-apiofuranose, the so-called apiogalacturonans (AGA)[36,37] and/or HGA that is substituted at the 3 position with D-xylose[32-35], so-called xylogalacturonan (XGA). RG-II is a complex polysaccharide that accounts for approximately 10-11% of pectin[8,39]. RG-II has an HGA backbone with four structurally complex side chains attached to C-2 and/or C-3 of the GalA[8,27] (Fig. 3). Rhamnogalacturonan I (RG-I) accounts for 20-35% of pectin[39] (Fig. 3). RG-I is a family of polysaccharides with an alternating [→4)-α-D-GalA-(1→2)- α-L-Rha-(1→] backbone in which roughly 20-80% of the rhamnoses are substituted by arabinan, galactan, or arabinogalactan side branches[6,8,30].

Pectins are believed to have multiple roles during plant growth, development, and in plant defense responses. For example, pectic polysaccharides play essential roles in cell wall structure[43], cell adhesion[44] and cell signaling[45,46]. Pectins also appear to mediate pollen tube growth[47] and to have roles during seed hydration[48,49], leaf abscission[50], water movement[51], and fruit development[47,8]. Oligosaccharides cleaved from pectin also serve as signals to induce plant defense responses[52,53]. Studies of mutant plants with altered wall pectin reveal that modifications of pectin structure leads to dwarfed plants[43], brittle leaves[44], reduced numbers of side shoots and flowers[54], malformed stomata[44] and reduced cell adhesion[55].

Although pectins appear to have multiple roles in plants, in no case has their specific mechanism of action been determined. One way to directly test the biological roles of pectins, and to study their mechanisms of action, is to produce plants with specific alterations in pectin structure. This can be done by knocking out genes that encode the pectin biosynthetic enzymes. Such enzymes include the nucleotide-sugar biosynthetic enzymes and the glycosyltransferases that synthesis the pectic polysaccharides. Each glycosyltransferase is expected to transfer a unique glycosyl residue in a specific linkage onto a specific polymeric/oligomeric acceptor. To date, only five[56-59] of the more than 200 predicted wall biosynthetic glycosyltransferases have been identified at the gene level, and none of these have been shown to encode pectin biosynthetic enzymes. The progress report below describes our identification of the first gene (*GALAT1*) encoding a pectin biosynthetic enzyme, GALAT1. The identification of this gene now allows the characterization of this biosynthetic enzyme, the use of this gene to produce mutated enzyme and plants, and studies of the role of this specific GalAT in homogalacturonan

and pectin synthesis. For clarity, please note that in this proposal we refer to the gene as *GALAT1*, to the protein produced by the gene as GALAT1, and to an α1,4-galacturonosyltransferase enzyme activity in plant extracts as GalAT.

Pectin Synthesis

Based on the known structure of pectin, at least 58 distinct glycosyl-, methyl- and acetyl-transferases are required to synthesize the family of polymers known as pectin. As shown in the review[9] in Appendix I and in Table I, between 4-9 galacturonosyltranferases are predicted to be required for the synthesis of HGA, RG-I, RG-II and possibly for the synthesis of the modified forms of HGA known as XGA and AGA. Since, as described in the progress report below, the *GALTA1* gene was identified based on the presence of the GALAT1 protein in *Arabidopsis* partially purified protein fractions that contained UDP-GalA:HGA α1,4GalAT enzyme activity, we propose that GALAT1 is UDP-GalA:HGA α1,4GalAT. Thus, the following review of pectin synthesis will concentrate on prior research on GalAT. Additional information on other pectin biosynthetic enzyme activities can be found in the review in Appendix 1[9].

UDP-GalA:Homogalacturonan α-1,4-Galacturonosyltransferase (GALAT1)

Membrane-bound α1-4galacturonosyltransferase (GalAT) activity has been identified and partially characterized in mung bean[10,11], tomato[12], turnip[12], sycamore[13], tobacco suspension[2], radish roots[5], enriched Golgi from pea[7], Azuki bean[14], *Petunia*[15], and *Arabidopsis* (Sterling and Mohnen, unpublished results) (see Table II). The pea GalAT was localized to the Golgi[7] with its catalytic site facing the lumenal side of the Golgi,[7]. These results provide the first direct enzymatic evidence that the synthesis of HGA occurs in the Golgi. In *in vitro* reactions, GalAT adds [^{14}C]GalA from UDP-[^{14}C]GalA[1,60] onto endogenous acceptors in microsomal membrane preparations to produce radiolabeled products of large molecular mass (i.e. ~105 kd in tobacco microsomal membranes[2] and ≥ 500 kd in pea Golgi[7]). The cleavage of up to 89% of the radiolabeled product into GalA, digalacturonic acid (diGalA) and trigalacturonic acid (triGalA) following exhaustive hydrolysis with a purified endopolygalacturonase confirmed that the product.

Table II. *Comparison of apparent catalytic constants and pH optimum of HGA-α1,4-galacturonosyltransferases*[1,2]

Enzyme[2]	Plant Source	Apparent K_m for UDP-GalA (μM)	pH optimum	Vmax (pmol $mg^{-1}min^{-1}$)	Ref
GalAT[1]	mung bean	1.7	6.0	~4700	10
GalAT	mung bean	n.d.	n.d.	n.d.	61
GalAT	pea	n.d.[5]	6.0	n.d.	62
GalAT	pea	n.d.	n.d.	n.d.	7
GalAT	sycamore	770	n.d.	?	13
GalAT	tobacco	8.9	7.8	150	2
GalAT (sol)[3]	tobacco	37	6.3-7.8	290	3
GalAT (sol)[3]	Petunia	170	7.0	480	15
GalAT (per)[4]	Azuki bean	140	6.8-7.8	2700	14

[1] Adapted from[6]

[2] Unless indicated, all enzymes are measured in particulate preparations

[3] (sol): detergent-solubilized enzyme

[4] (per): detergent-permeabilized enzyme
[5] n.d.: not determined

synthesized by tobacco GalAT was largely HGA. Thus, *in vitro* the crude enzyme catalyzes the reaction: UDP-GalAT + HGA(n) → HGA(n+1) + UDP. The product produced *in vitro* in tobacco microsomes was ~ 50% esterified[2] while the product produce in pea Golgi did not appear to be heavily esterified[7]. These results suggest that the degree of methyl esterification of newly synthesized HGA may be species specific and that methylesterification occurs after the synthesis of at least a short stretch of HGA. GalAT in detergent-permeabilized microsomes from azuki bean seedlings added [^{14}C]GalA from UDP-[^{14}C]GalA onto acid-soluble polygalacturonate (PGA) exogenous acceptors[14]. Treatment of the radiolabeled product with a purified fungal endopolygalacturonase yielded GalA and diGalA, confirming that the activity identified was a GalAT comparable to that studied in tobacco and pea. The azuki bean enzyme had a surprisingly high specific activity of 1300-2000 pmol mg^{-1} min^{-1}, especially considering the large amount (3.1-4.1 nmol mg^{-1} min^{-1}) of polygalacturonase activity that was also present in the microsomal preparations. As with the product made by tobacco, no evidence for the processive transfer of galactosyluronic acid residues onto the acceptor was obtained (see below).

GalAT can be solubilized from membranes with detergent[3]. Solubilized GalAT adds GalA onto the non-reducing end[4] of exogenous HGA (oligogalacturonide; OGA) acceptors of a degrees of polymerization ≥ ten[2]. The bulk of the HGA elongated *in vitro* by solubilized GalAT from tobacco membranes[3], or detergent-permeabilized Golgi from pea[7], at roughly equimolar UDP-GalA:acceptor concentrations is elongated by a single GalA residue. These results suggest that solubilized GalAT *in vitro* acts nonprocessively, (i.e. distributively). The apparent lack of *in vitro* processivity of GalAT was recently confirmed by Akita *et al.* who, using pyridylaminated oligogalacturonates as substrates and high concentrations of UDP-GalA, showed that although OGAs can be elongated in a "successive" fashion with up to 10 GalA residues by solubilized enzyme from petunia pollen[15], the kinetics of this response suggest a distributive mode of action. Our working hypothesis is that for a Golgi-localized enzyme that synthesizes a complex polymer in a confined internal cellular compartment with sufficiently high concentrations of substrate, it is would not necessarily be advantageous for the enzyme to act processively. In fact, the reaction velocity could be hindered under such conditions if the enzyme were processive[65].

The apparent kinetic constants and pH optimum for the characterized GalATs are shown in Table II. We have performed additional non-published kinetic studies in tobacco and radish that suggest that solubilized and membrane bound GalAT may have unusual apparent biphasic kinetics. We tested Vo for radish GalAT at 2 µM to 80 mM UDP-GalA and obtained a biphasic curve (Fig. 4), suggesting that the kinetics of GalAT, at least in the membrane and soluble fractions, are complex. Comparable results were also obtained for the solubilized radish and tobacco enzyme. The initial Vo vs [UDP-

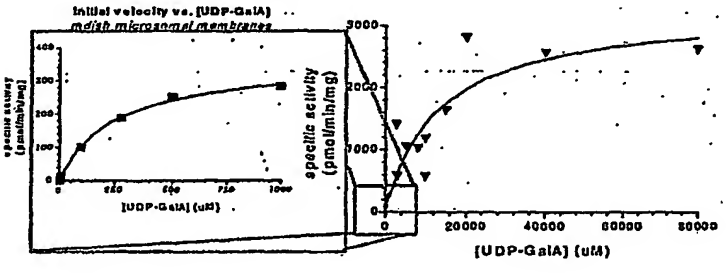


Fig. 4 ***GalAT kinetics in radish microsomal membranes.*** **Radish microsomal membranes (60 - 80 µg protein) were incubated with 70 µg OGA (DP 7 - 23) and the indicated concentrations of UDP-GalA. Each reaction contained a small concentration of UDP-[^{14}C]GalA (2 - 3.6 µM) with larger amounts of nonradioactive UDP-GalA. The precipitated reaction products were measured by liquid scintillation counting. The data are the averages of duplicate samples from three separate experiments. The Y axis is specific activity (pmol min^{-1} mg^{-1}).**

GalA] curve was hyperbolic and appeared to reach an initial maximum Vo of ~ 300 pmol mg^{-1} min^{-1} at ~1 mM UDP-GalA, confirming previous results reported for tobacco[2,3]. However, at ≥ 2 mM UDP-GalA there was a second hyperbolic increase in GalAT activity that reached a maximum of ~2-4 nmol min^{-1} mg^{-1} with ~20 mM UDP-GalA. In crude enzyme preparations it was not possible to determine the basis for the unusual kinetics. One possibility is that two GalATs were present, one with a low Km and one with a high Km. Another possibility is that UDP-GalA is both a substrate and an allosteric regulator of GalAT. Alternatively, a more "trivial" explanation is that at low substrate concentrations the kinetics of GalAT were effected by a catabolic enzyme (e.g. a phosphodiesterase) in the enzyme prep. The identification of the *GALAT1* gene (see progress report) and the production of pure heterologously expressed GALAT1 now allow us to study this further.

PROGRESS REPORT

Identification of the GALAT1 Gene (results from current USDA grant)

The major goal of our current USDA grant, in which we are in the beginning of the second year, was to identify the *Arabidopsis* gene for GalAT by heterologous functional expression. As outlined below, we have been successful in transiently expressing GALAT1 in mammalian cells and identified the first gene for a GalAT. We are currently characterizing the expressed GalAT (a third goal of the current funded grant) and are in the process of making stable mammalian cell transformants. Due to our success with the heterologous expression, and to the lack of available T-DNA mutant lines in our identified GalAT gene, we have not proceeded with the goal of identifying GalAT mutant plants. We postponed the RNAi studies due to the reduction in the requested funding of our last grant (see rather Specific Aim 4 in this renewal proposal).

We used a partial purification-tandem mass spectrometry approach to identify putative GalAT genes from *Arabidopsis* (see Fig. 5 for schematic of strategy). GalAT from *Arabidopsis* was partially purified from detergent-solubilized enzyme by sequential passage over two or more of the following resins: cation exchange resin SP-Sepharose, reactive green 19 resin, reactive

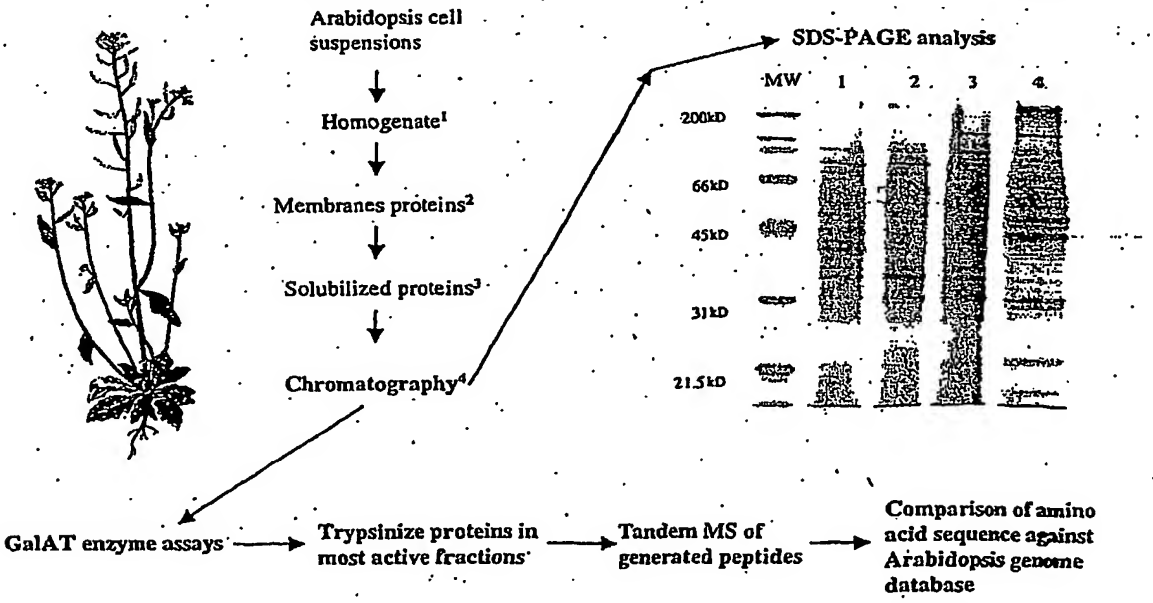


Figure 5. *Outline of strategy to identify the gene for GalAT.* The sequenced *Arabidopsis* genome allowed the use of a function-based partial purification-mass spectrometry approach to identify the putative carbohydrate biosynthetic genes. 1: homogenate, 2: total membranes, 3: solubilized proteins, 4: initial anion exchange purification step, 5: additional chromatography (not shown).

blue 72 resin, reactive yellow 3 resin, and UDP-agarose. Proteins obtained from selected fractions from these columns were treated with trypsin to generate peptides, and the amino acid sequence of the peptides identified by liquid chromatography-tandem mass spectrometry. The amino sequence was used to screen the *Arabidopsis* gene/protein database. Thirty unique proteins were solely identified in the GalAT-containing fractions (i.e. not present in fractions not containing GalAT activity). Among the 30 unique proteins that co-purified with GalAT activity, two proteins (designated JS33 and JS36) were identified as *Arabidopsis* putative GalAT proteins/genes based on their having at least one predicted transmembrane domain and since they contained a predicted glycosyltransferase domain (see CAZy database; http://afmb.cnrs-mrs.fr/CAZY/index.html).

These two genes, along with another *Arabidopsis* gene with high sequence similarity to JS36 (designated JS36L for JS36-like) (see below) were either cloned by RT-PCR (JS36 and JS36L) using mRNA from *Arabidopsis* flower and stem tissue (gift of Maor Bar-Peled, CCRC), or a cDNA clone was obtained from the Arabidopsis Biological Resource Center (JS33). The proteins encoded by these genes each have a predicted single transmembrane domain (Table III). The genes were truncated to remove their N-terminal region including all or most of the predicted transmembrane domain (see Table III), and the truncated genes were inserted into the mammalian expression vector pEAK (gift of Kelley Moremen lab, CCRC).

Table III. ***Predicted characteristics of JS36, JS33 and JS36L proteins.*** Predictions made using information from the NCBI database and the SOSUI (Classic & Membrane Prediction program) at BCM Search Launcher site (http://searchlauncher.bcm.tmc.edu/seq-search/struc-predict.html).

Gene	NCBI protein ID	# amino acids	MW (kd)	pI	Predicted transmem-brane domain	Truncated protein
At3g61130 (JS36)	NP_191672	673	77.4	9.95	$^{N}22\text{-}44^{C}$	$^{N}42\text{-}673^{C}$
At2g38650 (JS33)	NP_565893	619	69.7	8.63	$^{N}23\text{-}45^{C}$	$^{N}44\text{-}619^{C}$
At5g47780 (JS36-like)	NP_568688	616	71.1	9.26	$^{N}6\text{-}22^{C}$	$^{N}26\text{-}616^{C}$

The truncated forms of JS33, JS36 and JS36L in a pEAK vector (Edge BioSystems as modified by Moremen lab, CCRC) containing a heterologous signal sequence, a histidine (HIS) tag, and two influenza hemagglutenin (HA) epitopes, were transiently expressed in human embryonic kidney cells (HEK293 cells) for 46 hours. The media was collected. Since the translational fusion proteins constructed contained two copies of the HA epitope, the media surrounding the transiently infected cells was incubated with a mouse anti-HA IgG1 that was bound to Protein A Sepharose and the immunoabsorbed protein was assayed for GalAT activity. Table IV-A shows that the JS36 construct expressed GalAT activity. The experiment was repeated using Protein G Sepharose (Table IV-B) as an immunoabsorbent for the mouse anti-HA IgG1 (since Protein G Sepharose is reported to bind IgG1 more effectively than Protein A Sepharose). The results show that JS36 again expressed GalAT activity, although the activity was lower, suggesting that the enzyme was losing activity over time. From these results of the transiently expressed protein we conclude that JS36 is a GalAT and have named the gene *GALAT1* in accordance with the rules put forward by the "Commission on Plant Gene Nomenclature (http://genome-www.stanford.edu/Mendel/) and The Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org/nomencl.html; http://www.arabidopsis.org/info/guidelines.html#genes).

Table IV. Anti-HA epitope:protein A Sepharose immunoabsorbed protein from media surrounding mammalian HEK293 cells transiently expressing recombinant putative GalAT. Truncated JS36, JS33 and JS36L genes inserted into the mammalian expression vector pEAK, along with a pEAK vector only control, were independently transfected into human embryonic kidney cells (HEK293 cells). The transfected cells were grown for 46 hours to transiently express the recombinant proteins and the media surrounding the cells was incubated with anti-HA IgG1 bound to either Protein A Sepharose (A) or Protein G Sepharose (B). The immunoabsorbed protein was assayed for GalAT activity using UDP-[^{14}C]GalA and oligogalacturonides as substrates. The data are the average [^{14}C]GalA incorporated into product (minus T0 controls) from duplicate 45 min (A) or triplicate 60 min (B) reactions. Time 0 controls are reactions to which base was added to inactivate the enzyme prior to the addition of substrates. The GalAT activity in solubilized Arabidopsis protein was assayed as a positive control.

Enzyme	(A) [^{14}C]GalA incorporated	(B) [^{14}C]GalA incorporated
Recombinant truncated JS36	161	112
Recombinant truncated JS33	30	0
Recombinant truncated JS36L	0	0
Vector control	0	4
Solubilized Arabidopsis enzyme (control)	1434	550

As mentioned above, analysis of the gene sequence of GALAT1 shows that the predicted expressed protein contains one transmembrane domain. This is in agreement with the GalAT activity being membrane bound in all species tested (see appendix[9]). Furthermore, the predicted topology of GALAT1 is that of a type-II membrane protein, in agreement with our previous determination that the catalytic site of pea GalAT lies in the lumen of the Golgi. Type-II membrane proteins have a short N-terminal cytosolic tail, a transmembrane region, a stem region, and a C-terminal catalytic domain[16].

GALAT1 is a member of the Glycosyltransferase Family 8 in the CAZy database [database of putative and proven carbohydrate modifying enzymes that currently contains 61 different proposed glycosyltransferase families (http://afmb.cnrs-mrs.fr/CAZY/index.html)[66,67]]. The presence of GALAT1 in Family 8 is in agreement with our demonstrated activity of GALAT1 as an α1,4-galacturonosyltransferase, since Family 8 is a family of proposed retaining glycosyltransferases and GALAT1 is a retaining enzyme (i.e. the α-configuration in the substrate UDP-α-GalA is retained in the product α1,4-linked-galacturononan (HGA).

Of the 194 putative glycosyltransferase proteins in Family 8, only two have determined 3D crystal structures: an α1,4-galactosyltransferase from *Neisseria meningitidis* and glycogenin. A pairwise blast shows that GALAT has no sequence similarity with glycogenin but does have sequence similarity with α-galactosyltransferase LgtC (22% identity and 43% similarity over a 189 amino acid region). The *Neisseria meningitidis* GalT LgtC adds α-galactose from UDP-Gal to a terminal lactose acceptor in a bacterial lipooligosaccharide acceptor to form an α1,4-linked galactose. Thus, the nucleotide-sugar substrates used by these two enzymes differ only in the UDP-Gal having a hydroxyl rather than a carboxylate at C6. The acceptors, however, are structurally different.

Interestingly, when we used the GALAT1 amino acid sequence to search for predicted protein structure using the GenTHREADER fold prediction program[68] on the PSIPRED protein

structure prediction server[69] (http://bioinf.cs.ucl.ac.uk/psipred/), the only crystallized protein to shown significant similarity was PDB entry 1GA8 which was the family 8 *Neisseria* GalT LgtC[70] mentioned above. GALAT1 is predicted to fold very similarly to the GalT LgtC. The large N-terminal domain of GalT LgtC (which is similar to the C-terminal domain of GALAT1 (i.e. the presumed catalytic domain) has a mixed α/β fold. Based on our comparison of the conserved amino acid sequences between the two proteins and the identity of specific amino acids that interact with the nucleotide-sugar and acceptor as determined by the crystal structure of GalT LgtC in complex with a UDP-Gal analog and the acceptor, we propose that in GALAT the amino acids D498, D500 and H635 are involved in complexing with Mn^{++}; while D499, G638, K641 are involved in binding to UDP; and R481 is involved in binding GalA in the acceptor. Thus, these amino acids will be among the first targeted for site directed mutagenesis (see below). It is also interesting to note that GalT LgtC binding of UDP-Gal appears to induce a conformational change in the protein, suggested induced fit. The nucleotide-sugar binds in a deep, solvent shielded cleft in GalT LgtC such that the UDP-Gal is almost entirely buried within the enzyme, possibly explaining the requirement for the nucleotide-sugar in order to crystallize GalT LgtC[70].

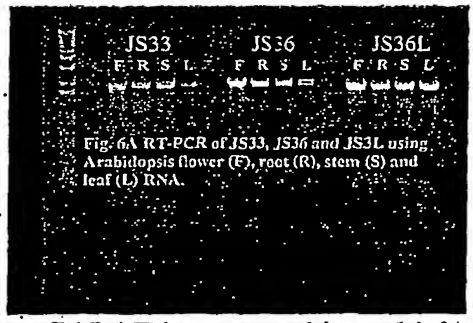


Fig. 6A RT-PCR of JS33, JS36 and JS3L using Arabidopsis flower (F), root (R), stem (S) and leaf (L) RNA.

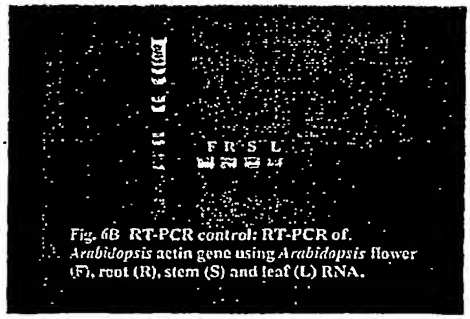


Fig. 6B RT-PCR control: RT-PCR of *Arabidopsis* actin gene using *Arabidopsis* flower (F), root (R), stem (S) and leaf (L) RNA.

GALAT is expressed in multiple *Arabidopsis* tissues at multiple times during development. We base this on our RT-PCR analysis of RNA from *Arabidopsis* flower, root, stem and leaf tissue (Fig. 6) showing that GALAT is expressed in all these tissues, and based on the 18 EST entries for this gene in the TAIR database (http://www.arabidopsis.org/) indicating that GALAT1 is expressed in developing seed, green siliques, roots and above ground organs.

Identification of a proposed GALAT1 Gene Family

A standard protein blast and a PSI Blast of the NCBI protein database using the GALAT1 (JS36) amino acid sequence reveals that GALAT1 is a member of a 15 member putative GalAT gene family in *Arabidopsis*. The genes we selected for this family had ≥ 30% amino acid identity and ≥ 50% amino acid similarity based on a PSI Blast. We further compared these genes along their entire coding sequences with JS36 using a Pairwise BLAST (Table VI) showing that this family of genes has ≥ 34% identity and ≥ 52% to JS36 in the portion of the genes C-terminal to the membrane spanning domain. This identity is comparable to the 37-54% identity shared among the proposed ten member *Arabidopsis* fucosyltransferase gene family (AtFU1-10)[71].

Of the members of the putative GalAT gene family, only GALAT1 has proven function. However, it is interesting that recently a mutant named *Quasimodo* was identified[55] that has 53% amino acid identity and 72% similarity to GALAT1. Thus, the gene affected in *Quasimodo* is a member of our proposed putative GalAT family. *Quasimodo* has 25% reduced levels of GalA in the cell wall and reduced cell adhesion, characteristics consistent with a GalAT. However, there is not yet enzyme confirmation of the activity of the *Quasimodo* gene.

The conserved amino acids in our proposed putative GalAT gene family are shown in Fig. 7. Glycosyltransferases are expected to contain one or more carboxylates at the catalytic site. At least one of the carboxylates is expected to coordinate a divalent cation associated with the nucleotide-sugar. In many glycosyltransferases the metal coordination involves two carboxylates that are often present as DDx, xDD, or DDD (the so-called "D(x) D" motif)[72]. We predict, based on a comparison with the crystal structure of the GalT LgtC (see previous section), and on the conservation within the putative GalAT gene family, that D498 and D500 are part of the "D(x) D' motif. However, additional structural and mutational data is required to confirm this.

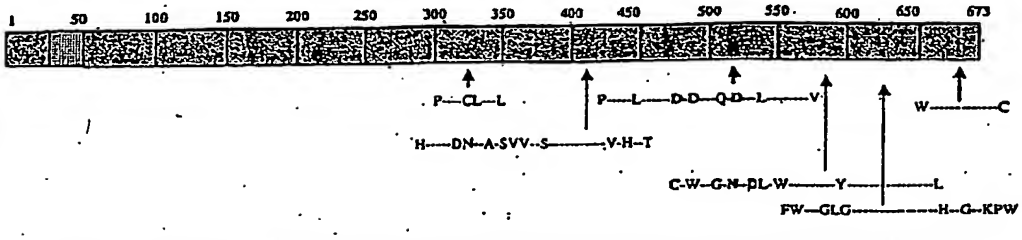


Figure 7. *Schematic representation of transmembrane spanning region and conserved amino acids in the Arabidopsis thaliana putative GalAT gene family.* The relative position of strictly conserved residues among the members of the putative GalAT family are numbered as for JS36 (i.e.GALAT1). The white speckled region from residues 22-44 represents the predicted transmembrane region.

Although it is predicted that retaining glycosyltransferases have a covalent mechanism for catalysis, in which the glycosyl residue is covalently transferred to the enzyme from the nucleotide-sugar prior to transfer to the oligo/polysaccharide acceptor[73], recent structural analysis of bovine (α-1,3-GalT) opens the possibility that, at least for bovine α-1,3-GalT, a sequential ordered binding mechanism may occur[74]. Clearly, additional structural data from other retaining glycosyltransferases is needed to determine the range of enzyme mechanisms used by retaining glycosyltransferases. If time allows, part of the proposed research is aimed at obtaining x-ray crystallographic and NMR 3D structural data on *Arabidopsis* GALAT1.

Table VI. Pairwise sequence alignment between JS36 and the other members of proposed GalAT gene family. The alignment was done using the NCBI Pairwise BLAST and Matrix Blosum62. The % amino acid identity and similarity are shown. In all cases the alignment compares the bulk of the C-terminal portion of the proteins on the carboxy-terminal side of the transmembrane region.

Gene	NCBI protein ID	%	% Similarity
**t3g61130 (*GALAT1*; JS36)	NP_191672	100%	100%
At5g47780 (JS36-like)	NP_568688	63%	81%
At2g46480	NP_182171	61%	75%
At4g38270	NP_195540	55%	73%
At3g25140 (*Quasimodo*)	NP_189150	53%	72%
At1g18580	AAK93644	48%	67%
At3g02350	NP_566170	47%	66%
At2g20810	NP_565485	46%	68%
At1g06780	NP_563771	44%	63%
At2g30575	NP_671901	43%	65%
At3g01040	NP_186753	42%	61%
At5g15470	NP_197051	42%	61%
At5g54690	NP_200280	38%	60%
At2g38650 (JS33)	NP_565893;	36%	60%
At3g58790	NP_191438	34%	52%

RATIONALE AND SIGNIFICANCE

Our identification of the GALAT1 gene in transiently transfected mammalian cells now allows the production of stably transformed cell lines that produce GALAT1 and experiments aimed at characterizing the mechanism of the enzyme and at determining the role of GALAT1 in pectin synthesis. The successful completion of the proposed research will provide the first biochemical characterization of a GalAT from any species. Specifically, the substrate specificity of GALAT1 will indicate whether it catalyzes only HGA synthesis, or also plays a role in RG-I and RG-II synthesis. Characterization of the kinetics of GALAT1 will clarify whether or not UDP-GalA is both a substrate and an allosteric regulator of the enzyme. Characterization of the mutated GALA1 enzyme will provide information regarding amino acids important in catalysis and substrate binding. The subcellular location of GALAT1 will provide the first framework for where, within the Golgi and plant endomembrane system the complex series of pectin biosynthetic reactions occur. If GALAT1 knockouts with a phenotype are obtained, the RNAi experiments will give information regarding the role of GALAT1 in pectin synthesis, potentially provide novel biosynthesis acceptors, and provide information about the role of pectin in plant growth and development. If time allows, this biosynthesis framework will be extended by the identification of GALAT1 binding proteins that would be putative pectin biosynthesis complex members. The successful completion of these studies will set the foundation for our future goals of a full *in vitro* reconstitution of functional pectin synthesis complexes. The proposed research will also compliment other research in the lab regarding the potential developmental and/or tissue specific expression of the *GALAT1* gene, at determining the *in vivo* role of GALTA1 via production of plants expressing altered level of GALAT1 transcript, and at determining the function of the other members of the proposed GALAT1 family.

GALAT1 has high sequence similarity to 14 other *Arabidopsis* proteins (the proposed putative GalAT family) and to proteins expressed in other plants. Possible GALAT1 homologs in other plants are a 68 kd protein expressed in *Cicer arietinum* (chickpea) epicotyls (76% amino acid identity; 87% similarity), a hypothetical protein from *Oryza sativa* (japonica) (59% identify; 75% similarity) and a protein from *Populus alba* (49% identity; 72% similarity). Thus, the results from the study of GALAT1 in *Arabidopsis* will likely extend to other plants, including those of high agricultural value.

EXPERIMENTAL PLAN

Specific Aim 1. Stable expression of GalAT1 in a heterologous expression system and detailed characterization of the enzyme.

Heterologous expression of GALAT1

In the progress report, we described results in which media from human embryonic kidney (HEK293) cells transiently infected with recombinant truncated *GALAT1* expressed GALAT. Whereas transient expression allowed the expression of sufficient GALAT to measure GalAT activity, additional expression strategies are required to produce large quantities of GALAT1 required for further characterization of the enzyme and for antibody production. Since the transiently expressed N-terminal epitope-tagged GALAT1 expressed in mammalian cells was active, one strategy will be to produce stable transfected clonal HEK293 lines[75] using the

described recombinant truncated GALAT1. The alternative strategy will be to express the full length and N-terminal truncated forms of GALAT1 in the fungal expression system *Pichia pastoris*. These systems were chosen since we and others[56-58] have successfully used them to express plant glycosyltransferases. Since it is not possible to predict, *a priori*, which of the systems will allow us to produce high levels of functional GALAT1, we will initially make constructs in both systems.

Kelley Moremen, a faculty at the CCRC routinely expresses proteins in mammalian cells and has advised us on the transient expression of GALAT1 in mammalian HEK293 cells. He will also advise us on the production and cloning of stable tranfected cell lines (see letter of support). The basic protocol will be as in the Edge BioSystems protocols with the modification suggested by the Moremen lab.

For expression in *P. pastoris*, cDNA encoding the entire, and the truncated soluble forms of, GalAT will be generated by PCR using gene/vector specific primers. The PCR products will be subcloned into appropriate *Pichia* expression vectors (Invitrogen)in which the cDNA is inserted downstream from an alcohol oxidase (AOX1) promoter. We are in the process of making full length coding sequence constructs for expression in the *Pichia* vector pPIC 3.5. This vector does not contain an epitope tag. We will also make epitope tagged *GALAT1* constructs in the *Pichia* vectors pPICz and pPICzα (Invitrogen) and determine whether functional C-terminal epitope-tagged constructs that do not effect GalAT activity can be recovered. We have already cloned the full coding sequence of GALAT1 into the plasmid vector pCR2.1 and are presently inserting the gene into the pPIC3.5 vector. *Pichia* will be transformed with the linearized recombinant construct. Abundant expression of the recombinant protein will be induced by growth in methanol. Cells will be plated and screened for growth under selection using an appropriate construct/selection scheme. The selected colonies will be assayed for GalAT activity. Approximately 10-20 GalAT positive colonies will be quantitatively assayed using the radioactive filter GalAT assay[63]. Advantages of the *Pichia* expression system include high levels of protein expression (up to 12 gm L^{-1}), the relative low cost of expression in *Pichia* compared to expression in mammalian systems, and the demonstrated success of *Pichia* for expression of glycosyltransferases[76-79]. Several labs at the CCRC are also currently using the *Pichia* system[80-82] and have successfully expressed plant enzymes. Once a high-GALAT1-producing line is recovered, production of large amounts of protein will be carried out in fermentors or spinner flasks (available at the CCRC and at the University of Georgia Fermentation Research Facility).

Characterization of Expressed GALAT1

To begin to address how HGA is synthesized, the kinetics, substrate specificity, and structure of the purified and recombinantly expressed GALAT1 will be determined and compared to the solubilized membrane-bound *Arabidopsis* GALAT purified by immunoabsorption using the polyclonal-antiGALAT1 (see below). Although the characteristics of GALAT1 are consistent with the enzyme being the/a catalytic subunit of the HGA synthase, GALAT1 could be a GalAT involved in RG-II or RG-I synthesis. For example, GalAT could represent an RG-I:GalAT that initially elongates HGA by a single GalA and then waits for a required NDP-Rha to start RG-I backbone synthesis. The kinetics of purified and recombinantly expressed GALAT1 for UDP-GalA and a size range of homogalacturonan and pectin acceptors will be determined. The effect of other nucleotide-sugars and oligosaccharide substrates on GalAT will also tested to identify activators and inhibitors. Our results with tobacco GalAT

indicate that UDP, UDP-GlcA, UDP-Gal, and UDP-β-GalA are not substrates but rather inhibit GalAT. We will confirm these observations with the recombinantly expressed GALAT1 and also test UDP-Xyl (substrate for xylogalacturonan) and UDP-Rha[83] (substrate for RG-I backbone) (gift of Maor Bar-Peled, CCRC).

The expressed full length and truncated enzymes will be assayed in reaction buffer in the presence, and absence, respectively, of Triton X-100. The kinetics of the enzyme for UDP-GalA will be carried out in a total of 1 μM to 80 mM UDP-GalA + UDP-[^{14}C]GalA. We routinely synthesize UDP-[^{14}C]GalA either by the 4-epimerization of UDP-[^{14}C]GlcA[1] or oxidation of UDP-[^{14}C]Gal[84] since UDP-[^{14}C]GalA is not commercially available. The effect of different acceptors on GALAT1 activity will be conducted using 100 μM UDP-GalA and 0.1-100 μg acceptor/ 30 μl reaction. The acceptors that will be tested include HGA oligosaccharides (oligogalacturonides) of degrees of polymerization ranging from 2-16, polygalacturonic acid, commercially available citrus pectin of ~30, 60 and 90% esterification, RG-I and RG-II. The products made using the different acceptors will be characterized[23]. If RG-I is shown to serve as an acceptor, RG-I backbone fragments that have a GalA or a Rha at the non-reducing end (gifts of Tadashi Ishii and Vons Foragen) will be used to determine acceptor specificity. The acceptors will be tested using both the precipitation assay[2] and the filter assay[63]. The enzymes will also be characterized for pH optimum, temperature sensitivity, and effect of reducing agents, divalent cations and salts. Finally, we have frequently observed an exopolygalacturonase activity in some of our most purified GalAT preparations from *Arabidopsis*. Exopolygalacturonase has also been reported in GalAT preparations from petunia[15] and Azuki bean[14]. We will assay expressed GALAT1 using a nonradioactive GalAT-PAGE assay[85] to determine if the expressed enzyme has any exopolygalacturonase activity.

Characteristics of the recombinant truncated GALAT1 will be compared to the GALAT1 solubilized from *Arabidopsis* membranes by immunoabsorption of the solubilized GALAT1 using anti-GALAT1 antibody (see section below) bound to Protein A or G Sepharose, or by coupling the anti-GALAT1 antibodies to 3M-Emphaze resin[86] and using the resin used to purify GALAT1 from solubilized *Arabidopsis* enzyme. If the characteristics of the immunoabsorbed *Arabidopsis* GALAT1 are different from the recombinant truncated GALAT1, we will analyze the immunoabsorbed GALAT1 by LC tandem mass spectrometry to determine if additional proteins were immunoabsorbed with the *Arabidopsis* solubilized GALAT1 that may have modified the activity (e.g. a heteromeric complex).

The recombinant GALAT1 and the GALAT1 immunoabsorbed-from *Arabidopsis* solubilized membranes will also by treated with *N*-glycanase to determine if they are *N*-glycosylated. To determine if they are O-glycosylated, the proteins will be exhaustively treated with *N*-glycanase, the released oligosaccharides removed by passage of the treated protein over a C18 SEP-PAK cartridge, and the resulting protein analyzed by TMS methylation analysis to determine the glycosyl residue composition of any carbohydrates still attached to the protein. Any oligosaccharide released by the N-glycanase treatment would also be analysis by TMS methylation. The results of these experiments would indicate whether the native *Arabidopsis* GalAT is glycosylated and whether the recombinant forms have the same or different glycosylation pattern. Changes in glycosylation could effect GALAT1 enzyme activity and/or substrate binding. GALAT1 is predicted to have 5 or 6 *N*-glycosylation sites (NetNGlyc 1.0 Prediction; http://www.expasy.org/sitemap.html).

As mentioned above, we have found that membrane-bound and solubilized GalAT activity in tobacco and radish has unusual apparent biphasic kinetics. Thus, we are particularly interested in determining if the expressed GALAT1 shows the same kinetics, including possible allosteric regulation by UDP-GalA. Such results would indicate likely 4° structure. We will test for possible multimeric structure by determining the mass of the enzyme by size exclusion chromatography and comparing these with the mass obtained by SDS-PAGE. The possibility that GALAT1 exists as a heteromultimer will be tested by mixing expressed GALAT1 with solubilized *Arabidopsis* enzymes and immunoabsorbing GALAT1 and proteins bound to it using either an anti-GALAT1 antibody or an anti-HA epitope antibody (see previous section).

Specific Aim 2. Production of a series of mutated GalAT1 proteins by site-directed mutagenesis of the *GALAT1* gene and heterologous expression of the mutated proteins.

The long term goal of the research proposed in this section is to determine the role, in GalAT1, of the 45 strictly conserved amino acids among the 15 members of the putative GalAT family. However, in the present proposal we will initially target, by site-directed mutagenesis, 6-10 amino acids likely involved in substrate/acceptor binding and/or catalysis. We will proceed to the other amino acids as time allows. The strategy will be to systematically mutate selected residues in GALAT1 to determine their roles in GALAT1 activity. The effect of these mutations on GALAT1 specific activity, and where warranted, on Km, Vmax, and acceptor specificity (i.e. OGA, RG-I and RG-II) and product size (i.e. enzyme processivity) will be determined. As noted in the introduction, we already propose a function for some of the amino acid residues based on analogy to their functions in the GalT LgtC (the most similar glycosyltransferase for which a 3D structure with substrates bound is available). We propose that D498, D500 and H635 are involved in complexing with Mn^{++}; D499, G638, K641 are involved in binding to UDP; and R481 is involved in binding GalA in the acceptor. Like many glycosyltransferases, GalAT requires Mn^{++}[2,3]. Based on the GalT LgtC structure, we propose that Mn^{++} is coordinated by the two phosphates oxygens of the UDP as well as by side chain oxygen from a D498, two side chain oxygens from a D500 and the ring hydrogen of H635. The structurally conserved amino acids between GALAT1 and GalT LgtC are all invariant in the 15 members of the proposed GalAT family except for D449 (invariant in 13 out of 15 genes) and R481 (invariant in 14 out of 15 genes). We propose that these residues are involved in catalysis and/or binding to UDP-GalA and HGA. It is also possible that some of the residues are required for structural integrity of GALAT or for binding of allosteric effectors. The mutations we propose to test initially will include: D498N, D500N, H635A, D499H (two of the family members have a His at this site), K641A and R481S. We will also mutate D563N, L564G, W567A, K568Q and K569Q since we hypothesize that these amino acids may be involved in catalysis or acceptor binding based on their conservation and/or similarity within the proposed putative GalAT gene family.

Site directed mutagenesis of truncated, and possibly full length (if we can get the full length GALAT1 expressed in a functional form) GALAT1 will be at a specific at amino acid(s) using the Promega GeneEditor in vitro site-directed mutagenesis system. The GeneEditor system has the advantage over the Altered Sites II kit in that any vector containing the wild type TEM-1 beta lactamase gene (i.e. ampicillin resistance gene) can be used. Most cloning vectors, including those we are using for expression in mammalian and *Pichia* cells, have this antibiotic

resistance gene. The mutant oligonucleotides to be used will include single base substitutions but eventually may also include multiple nucleotide substitutions or deletions. The Promega GeneEditor system can be used for multiple substitutions, deletions and insertions.

Specific Aim 3. Establishment of the subcellular location of GALAT1 using anti-GALAT1 antibodies and GFP fusion proteins.

Production and use of antibodies

Anti-GalAT antibodies are necessary for the proposed immunocytochemistry experiments, to immuopurify solubilized GALAT1 from *Arabidopsis*, and to select proteins that potentially bind to GALAT1 and may function in pectin biosynthetic enzyme complexes. The goal is to generate antibodies against the heterologously expressed truncated GALAT1, although if we find that the full length GALAT1 gives significantly more or greater synthetic activity, we may make antibodies against the full length forms as well. In theory one can make either polyclonal or monoclonal antibodies. However, polyclonals have the advantage that one can expect multiple antibodies with multiple points of interaction against GALAT1. Such antibodies would be useful for a range of types of experiments, including the subcellular immunocytochemistry, immunoprecipitation/absorption, and enzyme activity inhibition studies. Thus, our initial strategy will be to generate polyclonal antibodies. A secondary strategy would be to generate monoclonal antibodies if the polyclonals do not function for the proposed experiments.

Polyclonal antibodies will be generated in two ways to enhance the likelihood of obtaining useful GALAT1 specific antibodies. Our main strategy will be to use purified recombinantly expressed GALAT1 directly to generate rabbit polyclonal antibodies. While this should supply polyclonal antibodies that will recognize GALAT1, the possibility exists that, due to the sequence similarity of GALAT1 with other putative GalATs in our proposed putative GalAT gene family, the antibodies may also cross react with other members of the proposed gene family. Thus, we will also immunize rabbits with an eight-branched, multiple antigenic protein (MAP)[87] synthesized based on an N-terminal less-well conserved region of GALAT1 (i.e. a 14-15 non-conserved amino acid region N-terminal to the conserved C-terminal region). Colleagues at the CCRC have successfully used this approach to make anti-*Cochliobolus* specific anti-endopolygalacturonase antibodies[88]. While the strength of the MAP system is its ability to generate specific anti-peptide antibodies, the ability of the MAP generated antibodies to cross-react with the entire GALAT1 protein may be more limited. Thus, it may be necessary to make multiple MAPs based on multiple non-conserved regions. The MAP will be synthesized at the Molecular Genetics Instrumentation Facility, University of Georgia, using the 8-branch Fmoc resin as described[88]. The polyclonal antibodies will be raised at the University of Georgia Polyclonal Antibody Production Service using GALAT1 that we purify and provide. Monoclonal antibodies may be generated in mice and/or rabbits as a secondary strategy if necessary[89]. Dr. Michael Hahn has extensive experience in the generation and screening of antibodies and has agreed to advise us (see attached letter). The generation of murine hybridomas will be carried out at the Monoclonal Antibody Facility located in the School of Veterinary Medicine at the University of Georgia.

Polyclonal serum or hybridomas that secret antibodies against GalAT will be identified using a two-step procedure. The initial screen will be an enzyme-linked immunosorbent assay[86] and western blotting against purified recombinant GALAT1. The specificity of the antibodies

will be demonstrated by Western-blot analysis of expressed GALAT1 and of solubilized proteins from *Arabidopsis* cell suspensions. Positive anti-GALAT1 antibodies will also be subjected to second round screening of the serum or hydridoma supernatants with GALAT1 followed by immunoabsorption using Protein A-Sepharose[86,90]. Anti-GalAT antibodies will be identified by a loss of GALAT1 protein and GalAT activity in the supernatant and an elevation of GALAT1 protein and GalAT activity in the pellet compared to preimmune controls. If monoclonal antibodies are used, anti-GalAT-producing hybridomas will be subcloned and antibodies produced in ascites culture. The antibodies will be purified on protein G-Sepharose, and isotyped using a mouse monoclonal antibody isotyping kit (Sigma).

Subcellular localization of GALAT1

As mentioned in the introduction, all available data, including the localization of the catalytic domain of GalAT in the Golgi lumen, suggest that pectin is synthesized in the Golgi and transferred via vesicles to the wall. However, what is not known is how the different glycosyltransferases function to make specific pectin structures. Our expectation is that different glycosyltransferases are localized in a sequential manner to different cisternae of the Golgi[22,91] in an order indicative of the order in which pectin is synthesized as it moves from the cis, through the medial and to the trans Golgi. Evidence from both animal[92,93] and plants[94] suggests that, either individually or in combination, the transmembrane domain (i.e. the bilayer thickness model[95]), the N- or C-terminal sequences flanking the transmembrane domain, and/or the lumenal domain (i.e. the 'kin recognition model'[96]) contribute to localization of proteins within the Golgi system. We propose to determine the subcellular localization of GALAT1 within the Golgi by immunocytochemistry in order to provide additional information on role of GALAT1 in pectin synthesis. For example, a location of GALAT1 in the cis and medial Golgi cisternae would be consistent with a function of GALAT1 in HGA synthesis, while a localization primarily in the late medial or trans Golgi would be more suggestive of a role in RG-I or possible RG-II synthesis. It should be noted that such subcompartment localization studies, while important and novel for the pectin biosynthetic enzymes, are also novel in any species since "the precise location of only a small number of the glycosyltransferase proteins within the Golgi apparatus have been determined"[93].

Anti-GALAT1 antibodies will be used to identify where in the Golgi GALAT1 is localized. Both developing *Arabidopsis* seedlings and growing suspension cultures will be used as a tissue source since both tissues include cells actively making wall and since Michael Hahn and Glenn Freshour at the CCRC & Plant Biology Department have already established procedures to obtain excellent immunocytochemistry results from these tissues. The tissues will be prepared in two ways. For chemical fixation, four day old *Arabidopsis* seedlings with roots, or exponentially growing *Arabidopsis* suspension cultures will be fixed with potassium phosphate buffer (KPB) containing 2.5% (v/v) glutaraldehyde, the root tissue or suspension cells removed and washed with KPB, put through a graded ethanol dehydration series, and infiltrated with LR White embedding resin as described[97]. Alternatively, tissue will be prepared by high-pressure freezing and freeze substitution[91,98]. Cells or seedling root tissue will be suspended in 25% dextran (a cryoprotectant) and concentrated[91,98]. The cells will be transferred to sample holders pretreated with lecithin/chloroform, allowed to dry, frozen under high pressure[98] using a Balzers HPM010 High Pressure Freezer (available at the University of Georgia Electron Microscopy facility, Department of Plant Biology) and stored in liquid nitrogen until use. Samples will be freeze-substituted by transfer to precooled cryo-substitution vials containing 2%

osmium tetroxide and 8% dimethoxypropane in anhydrous acetone with substitution for 2-3 days at –80°C in an acetone dry ice bath. The samples will be warmed slowly, the substitution media discarded, the samples rinsed with acetone, and the tissue infiltrated with LR White resin[91,97].

Thin sections (<100 nm) will be cut using an MT 6000-XL ultramicrotome, the sections collected on Formvar-coated, gilded copper slot grids and the dried sections immunolabeled following hydration as described[97]. The sections will be blocked for non-specific binding by incubation in 3% (w/v) nonfat dried milk in KPBS and the sections incubated in the primary GALAT1 antiserum diluted in KPBS or with a pre-immune control antibody. The sections will be rinsed with KPBS and labeled by incubation on droplets of secondary anti-rabbit IgG conjugated to 15-nm colloidal gold[97]. After immunolabelling the sections will be poststained with 4% uranylacetate and then lead citrate[97] and examined with a Zeiss EM 902A Electron microscope.

We will use the morphology definitions of Zhang and Staehelin to distinguish the cis from the medial side of the Golgi[91]. However, the cis versus the trans end of the Golgi stacks will also be distinguished by immunolabeling some sections using anti-clathrin antibodies that label preferentially the trans-Golgi network[99]. Different sized gold particles will be used to distinguish the different antibodies. Michael Hahn has extensive experience in the use of antibodies for immunocytochemistry and has agreed to advise us (see attached letter).

Specific Aim 4. Use RNAi to generate GALAT1 knockouts and characterize the phenotype of the mutated plants.

Double-stranded RNA-mediated interference (RNAi) is a method to study the function of genes in plants[100]. Transgenic plants harboring an RNAi construct often have reduced expression of the gene-specific mRNA. The resulting plants may display either complete gene silencing, thus having a knockout phenotype, or a partial "knockout" phenotype due to 'leaky' expression. The RNAi approach should allow the suppression of *GALAT1* expression and a reduction or loss of GALAT1. This will allow us to elucidate the function of GALAT in pectin synthesis and in the plant. We will also continually monitor the TAIR database to determine if any T-DNA insert lines for GALAT become available. If so, we will order the seed and characterize the mutants as described below.

To generate GALAT knockout phenotypes using RNAi gene silencing, complementary gene-specific coding regions unique to GALAT1 (i.e. not found in the other gene family members) will be generated by fusing one sense region (~200 nt) to its identical region in the antisense orientation using GUS as a linker. A 700-bp partial coding linker DNA derived from the GUS gene will be separate the two complementary *gene* fragments. The DNA construct (sense):GUS:(antisense) will be generated by PCR and ligated into a Bluescript-based vector between the 35S cauliflower mosaic virus promoter and the octopine synthase terminator. The resulting expression cassette will be ligated into a binary vector, the vector transformed into *Agrobacterium tumefaciens* by electroporation, and the transformed *A. tumefaciens* used to transform *Arabidopsis thaliana* ecotype Columbia by vacuum infiltration[101]. T1 seedlings will be selected on agar plates containing the appropriate antibiotic/herbicide and resistant seedlings will be transferred to soil. T2 resistant seed will be collected from independent lines. Maor Bar-Peled, a colleague at the CCRC, routinely uses these methods in his lab will advise us as necessary. Transgenic homozygous plants will be selected and characterized (assuming they are not lethal). Of particular importance regarding pectin synthesis, the cell walls will be isolated

and analyzed for glycosyl residue composition and linkage to provide information about the possible role of GALAT1 in pectin synthesis.

Future research. If time permits we will initiate two types of additional studies.

Identifying members of proposed HGA biosynthetic complexes. There is growing evidence that glycoconjugates are synthesized by complexes of glycosyltransferases and other types of proteins[102]. For example, ganglioside synthesis occurs via a tightly regulated formation of multiple glycosyltransferase complexes[102]. If time allows we will attempt to identify protein members of proposed biosynthetic complexes by immunoabsorbing such proteins bound to GALAT1 using anti-GALAT1 antibodies or anti-HA epitope antibodies. The immunoabsorbed proteins would be identified by SDS-PAGE, removed from the gel, and their amino acid sequence determined by LC-tandem mass spectrometry.

Determining the 3D structure of GALAT1. Studies aimed at determining the 3D structure of expressed recombinant truncated GALAT1, and full length GALAT1 would be initiated. A long-term goal of the research is to use GALAT1 (and other pectin biosynthetic enzymes) to synthesize pectins with specific structures. To do this an understanding of the 3D structure of GALAT1 and an identification of the amino acids important for nucleotide-sugar and acceptor binding and for catalysis is needed. With structural information, directed changes in GALAT1 gene/polypeptide sequence can be made with predicted effects on substrate and catalytic specificity. Two approaches will be used to gain information about GALAT1 3D structure: X-ray crystallography and NMR. Ideally, we would like to have structural data on both the active truncated recombinant GalATs and on the full length enzyme. Whereas success is probable with the truncated forms, 3D structural determination of membrane proteins is much more challenging, and NMR methods may be more useful with such proteins. The University of Georgia is a research core for the NIH-funded Southeast Collaboratory for Structural Genomics (SECSG) (http://www.secsg.org/) which is part of the developing Georgia Structural Genomics Center. This facility includes a robotic high throughput crystal screening (HTP-CS) system in the X-Ray Crystallography Core which is headed by Bi-Cheng Wang (PI of the SECSG grant). We will use the robotic system in collaboration with B.C. Wang to generate GALAT1 crystals and determine GALAT1 3D structure (see letter of support from B.C. Wang). We will also initiate studies in collaboration with Jim Prestegard (Co-PI of SECSG, CCRC) who heads the NMR Core of SECSG to attempt to use high-field NMR spectroscopy and computer based structure predictions to obtain NMR solution conformation data for GALAT1. This will involve producing appropriate recombinant isotopically labeled GALAT1 and obtaining residual dipolar coupling data from solutions of the protein in the presence of substrate/acceptors. Success with both studies will require producing sufficient amounts (5-10 mg) of full length and truncated GALAT1 for the NMR and crystallography approaches. These will be generated from the proposed research.

Results from Prior NRI Support

In work supported by our previous USDA-NRI grants [9/15/1994-12/31-00] we identified and characterized a membrane-bound and solubilized α-1,4-galacturonosyltransferase from tobacco, radish, pea and Arabidopsis[1-4,6-9,39,60,63,85,103,104] (see also progress report).

References

1. K. Liljebjelke, R. Adolphson, K. Baker, R.L. Doong, and D. Mohnen, Enzymatic synthesis and purification of uridine diphosphate [^{14}C]galacturonic acid: a substrate for pectin biosynthesis, Anal. Biochem. 225:296-304 (1995).
2. R.L. Doong, K. Liljebjelke, G. Fralish, A. Kumar, and D. Mohnen, Cell free synthesis of pectin: identification and partial characterization of polygalacturonate 4-α-galacturonosyltransferase and its products from membrane preparations of tobacco (*Nicotiana tabacum* L. cv samsun) cell suspension cultures, Plant Physiol. 109:141-152 (1995).
3. R.L. Doong and D. Mohnen, Solubilization and characterization of a galacturonosyltransferase that synthesizes the pectic polysaccharide homogalacturonan, The Plant Journal 13:363-374 (1998).
4. H.V. Scheller, R.L. Doong, B.L. Ridley, and D. Mohnen, Pectin biosynthesis: a solubilized galacturonosyltransferase from tobacco catalyzes the transfer of galacturonic acid from UDP-galacturonic acid onto the non-reducing end of homogalacturonan, Planta 207:512-517 (1999).
5. D. Mohnen, H.F. Quigley, K.L. Adams, J.A. Lemons, R.L. Doong, F. Goubet, and J. Sterling, A multi-enzyme approach to study pectin biosynthesis, Annual Meeting of the American Society of Plant Physiology, July 24-July 28 Abst. No. 203:65(1999).(Abstract)
6. D. Mohnen, Biosynthesis of pectins and galactomannans, in: "Comprehensive Natural Products Chemistry, Vol. 3, Carbohydrates and Their Derivatives including Tannins, Cellulose, and Related Lignins", B.M. Pinto., ed., Elsevier, Oxford, pp. 497-527 (1999).
7. J. Sterling, H.F. Quigley, A. Orellana, and D. Mohnen, The catalytic site of the pectin biosynthetic enzyme α-1,4-galacturonosyltransferase (GalAT) is located in the lumen of the Golgi, Plant Physiol. 127:360-371 (2001).
8. B.L. Ridley, M.A. O'Neill, and D. Mohnen, Pectins: structure, biosynthesis, and oligogalacturonide-related signaling, Phytochemistry 57:929-967 (2001).
9. D. Mohnen, Biosynthesis of pectins, in: "Pectins and their Manipulation", G.B. Seymour et al., Blackwell Publishing and CRC Press, Oxford, pp. 52-98 (2002).
10. C.L. Villemez, A.L. Swanson, and W.Z. Hassid, Properties of a polygalacturonic acid-synthesizing enzyme system from *Phaseolus aureus* seedlings. Arch. Biochem. Biophys. 116:446-452 (1966).
11. H. Kauss and A.L. Swanson, Cooperation of enzymes responsible for polymerization and methylation in pectin biosynthesis. Z. Naturforsch. 24:28-33 (1969).
12. T.-Y. Lin, A.D. Elbein, and J.C. Su, Substrate specificity in pectin synthesis. Biochem. Biophys. Res. Commun. 22:650-657 (1966).
13. G.P. Bolwell, G. Dalessandro, and D.H. Northcote, Decrease of polygalacturonic acid synthase during xylem differentiation in sycamore. Phytochemistry 24:699-702 (1985).
14. Y. Takeuchi and Y. Tsumuraya, In vitro biosynthesis of homogalacturonan by a membrane-bound galacturonosyltransferase from epicotyls of azuki bean, Biosci. Biotech. Biochem. 65:1519-1527 (2001).
15. K. Akita, T. Ishimizu, T. Tsukamoto, T. Ando, and S. Hase, Successive glycosyltransfer activity and enzymatic characterization of pectic polygalacturonate 4-α-galacturonosyltransferase solubilized from pollen tubes of *Petunia axillaris* using pyridylaminated oligogalacturonates as substrates, Plant Physiol. 130:374-379 (2002).

16. R.A.F. Reithmeier and C.M. Deber, Intrinsic membrane protein structure: principles and prediction, in: "The Structure of Biological Membranes", P. Yeagle., ed., CRC Press, Boca Raton, pp. 337-393 (1992).
17. D.H. Northcote, The Golgi apparatus. Endeavor 30:26-33 (1971).
18. D.H. Northcote and J.D. Pickett-Heaps, A function of the Golgi Apparatus in polysaccharide synthesis and transport in the root-cap cells of wheat, Biochem. J. 98:159-167 (1966).
19. P.J. Harris and D.H. Northcote, Polysaccharide formation in plant golgi bodies. Biochim. Biophys. Acta 237:56-64 (1971).
20. R.W. Stoddart and D.H. Northcote, Metabolic relationships of the isolated fractions of the pectic substances of actively growing sycamore cells. Biochem. J. 105:45-59 (1967).
21. P.J. Moore, K.M.M. Swords, M.A. Lynch, and L.A. Staehelin, Spatial organization of the assembly pathways of glycoproteins and complex polysaccharides in the golgi apparatus of plants. J. Cell Biol. 112:589-602 (1991).
22. L.A. Staehelin and I. Moore, The plant Golgi apparatus: structure, functional organization and trafficking mechanisms, Annu. Rev. Plant Physiol. Plant Mol. Biol. 46:261-288 (1995).
23. W.G.T. Willats, C.G. Steele-king, L. McCartney, C. Orfila, S.E. Marcus, and J.P. Knox, Making and using antibody probes to study plant cell walls, Plant Physiol. Biochem. 38:27-36 (2000).
24. F. Goubet and D. Mohnen, Subcellular localization and topology of homogalacturonan methyltransferase in suspension-cultured *Nicotiano tabacum* cells, Planta 209:112-117 (1999).
25. M.P. Vannier, B. Thoiron, C. Morvan, and M. Demarty, Localization of methyltransferase activities throughout the endomembrane complex system of flax (*Linum usitatissimum* L) hypocotyls, Biochem. J. 286:863-868 (1992).
26. T. Bourlard, A. Schaumann-Gaudinet, M.-P. Bruyant-Vannier, and C. Morvan, Various pectin methyltransferase activities with affinity for low and highly methylated pectins, Plant Cell Physiol. 38:259-267 (1997).
27. M. O'Neill, P. Albersheim, and A. Darvill, The pectic polysaccharides of primary cell walls, in: "Methods in Plant Biochemistry, Volume 2", P.M. Dey., ed., Academic Press, London, pp. 415-441 (1990).
28. J.M. Lau, M. McNeil, A.G. Darvill, and P. Albersheim, Structure of the backbone of rhamnogalacturonan I, a pectic polysaccharide in the primary cell walls of plants, Carbohydr. Res. 137:111-125 (1985).
29. S. Eda, K. Miyabe, Y. Akiyama, A. Ohnishi, and K. Kato, A pectic polysaccharide from cell walls of tobacco (*Nicotiana tabacum*) mesophyll, Carbohydr. Res. 158:205-216 (1986).
30. N.C. Carpita and D.M. Gibeaut, Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth, Plant J. 3:1-30 (1993).
31. M.A. O'Neill, D. Warrenfeltz, K. Kates, P. Pellerin, T. Doco, A.G. Darvill, and P. Albersheim, Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester - in vitro conditions for the formation and hydrolysis of the dimer, J. Biol. Chem. 272:3869(1997).
32. H.A. Schols, E.J. Bakx, D. Schipper, and A.G.J. Voragen, A xylogalacturonan subunit present in the modified hairy regions of apple pectin, Carbohydr. Res. 279:265-279 (1995).

33. A. Kikuchi, Y. Edashige, T. Ishii, and S. Satoh, A xylogalacturonan whose level is dependent on the size of cell clusters is present in the pectin from cultured carrot cells, Planta 200:369-372 (1996).
34. L. Yu and A.J. Mort, Partial characterization of xylogalacturonans from cell walls of ripe watermelon fruit: Inhibition of endopolygalacturonase activity by xylosylation, in: "Pectins and Pectinases", J. Visser et al., Elsevier, Amsterdam, pp. 79-88 (1996).
35. G.O. Aspinall, Chemistry of cell wall polysaccharides. in: "The Biochemistry of Plants, Vol.3.", J. Preiss., ed., Academic Press, New York, pp. 473-500 (1980).
36. R.R. Watson and N.S. Orenstein, Chemistry and biochemistry of apiose, Adv. Carbohydr. Chem. Biochem. 31:135-184 (1975).
37. D.A. Hart and P.K. Kindel, Isolation and partial characterization of apiogalacturonans from the cell wall of *Lemna minor*, Biochem. J. 116:569-579 (1970).
38. H.A. Schols, M.A. Posthumus, and A.G.J. Voragen, Structural features of hairy regions of pectins isolated from apple juice produced by the liquefaction process, Carbohydr. Res. 206:117-129 (1990).
39. D. Mohnen, R.L. Doong, K. Liljebjelke, G. Fralish, and J. Chan, Cell free synthesis of the pectic polysaccharide homogalacturonan, in: "Pectins and Pectinases", J. Visser et al., Elsevier Science B.V. Amsterdam, pp. 109-126 (1996).
40. W.G.T. Willats, L. McCartney, W. Mackie, and J.P. Knox, Pectin: cell biology and propects for functional analysis, Plant Mol. Biol. 47:9-27 (2001).
41. J. An, M.A. O'Neill, P. Albersheim, and A.G. Darvill, Isolation and structural characterization of β-D-glucosyluronic acid and 4-*O*-methyl β-D-glucosyluronic acid-containing oligosaccharides from the cell-wall pectic polysaccharide, rhamnogalacturonan I, Carbohydr. Res. 252:235-243 (1994).
42. J. An, L. Zhang, M.A. O'Neill, P. Albersheim, and A.G. Darvill, Isolaton and structural characterization of endo-rhamnogalacturonase-generated fragments of the backbone of rhamnogalacturonan I, Carbohydr. Res. 264:83-96 (1994).
43. M.A. O'Neill, S. Eberhard, P. Albersheim, and A.G. Darvill, Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth, Science 846-849 (2001).
44. R.G. Atkinson, R. Schröder, I.C. Hallett, D. Cohen, and E.A. MacRae, Overexpression of polygalacturonase in transgenic apple trees leads to a range of novel phenotypes involving changes in cell adhesion, Plant Physiol. 129:122-133 (2002).
45. D. Mohnen and M.G. Hahn, Cell wall carbohydrates as signals in plants, Sem. Cell Biol. 4:93-102 (1993).
46. F. Côté, K.-S. Ham, M.G. Hahn, and C.W. Bergmann, Oligosaccharide elicitors in host-pathogen interactions generation, perception, and signal transduction, in: "Plant-Microbe Interactions", B.B. Biswas et al., Plenum Press, New York, pp. 385-432 (1998).
47. J.-C. Mollet, S.-Y. Park, E.A. Nothnagel, and E.M. Lord, A lily stylar pectin is necessary for pollen tube adhesion to an in vitro stylar matrix, Plant Cell 12:1737-1749 (2000).
48. T.L. Western, J. Burn, W.L. Tan, D.J. Skinner, L. Martin-McCaffrey, B.A. Moffatt, and G.W. Haughn, Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in Arabidopsis, Plant Physiol. 127:998-1011 (2001).
49. W.G.T. Willats, L. McCartney, and J.P. Knox, In-situ analysis of pectic polysaccharides in seed mucilage and at the root surface of *Arabidopsis thaliana*, Planta 213:37-44 (2001).

50. Z.H. González-Carranza, C.A. Whitelaw, R. Swarup, and J.A. Roberts, Temporal and spatial expression of a polygalacturonase during leaf and flower abscission in oilseed rape and Arabidopsis, Plant Physiol. 128:534-543 (2002).
51. K. Brown, Xylem may direct water where it's needed, Science 291:571-572 (2001).
52. S. Fry, S. Aldington, P. Hetherington, and J. Aitken, Oligosaccharides as Signals and Substrates in the Plant Cell Wall, Plant Physiol. 103:1-5 (1993).
53. N. Shibuya and E. Minami, Oligosaccharide signalling for defence responses in plant, Physiol. Mol. Plant Pathol. 59:223-233 (2001).
54. M. Skjot, M. Pauly, M.S. Bush, B. Borkhardt, M.C. McCann, and P. Ulvskov, Direct interference with rhamnogalacturonan I biosynthesis in Golgi vesicles, Plant Physiol. 129:95-102 (2002).
55. S. Bouton, E. Leboeuf, and G. Mouille, QUASIMODO1 encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in Arabidopsis, Plant Cell 14:2577-2590 (2002).
56. M.E. Edwards, C.A. Dickson, S. Chengappa, C. Sidebottom, M.J. Gidley, and J.S.G. Reid, Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis, Plant J. 19:691-697 (1999).
57. R.M. Perrin, A.E. DeRocher, M. Bar-Peled, W. Zeng, L. Norambuena, A. Orellana, N.V. Raikhel, and K. Keegstra, Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis, Science 284:1976-1979 (1999).
58. A. Faik, N.J. Price, N.V. Raikhel, and K. Keegstra, An *Arabidopsis* gene encoding an α-xylosyltransferase involved in xyloglucan biosynthesis, Proc. Natl. Acad. Sci. USA 99:7797-7802 (2002).
59. D.P. Delmer, A hot mutant for cellulose synthesis, Trends in Plant Science 3:164(1998).
60. A. Orellana and D. Mohnen, Enzymatic synthesis and purification of [^{3}H] uridine diphosphate galacturonic acid for use in studying Golgi-localized transporters, Analytical Biochemistry 272:224-231 (1999).
61. H.J. Crombie and J.S.G. Reid, A homogalacturonan synthase from mung bean hypocotyls, Cell Wall '01 - 9th International Cell Wall Meeting, September 2-7, 2001 Toulouse, France 131(2001).(Abstract)
62. C.M. Cumming and C.T. Brett, A galacturonyltransferase involved in pectin biosynthesis. in: "Cell Walls '86. Proceedings of the Fourth Cell Wall Meeting. Paris - September 10-12, 1986.", B. Vian et al., Université Pierre et Marie Curie - Ecole Normale Supérieure. Paris, pp. 360-363 (1986).
63. J.A. Lemons, I.F. Forkner, J. Sterling, and D. Mohnen, Development of a filter assay for measuring homogalacturonan: α1,4-galacturonosyltransferase activity (in preparation)
64. N. Williams, Rain forest fragments fare poorly, Science 278:1016(1997).
65. S. Pagès, H.C.M. Kester, J. Visser, and J.A.E. Benen, Changing a single amino acid residue switches processive and non-processive behavior of *Aspergillus niger* endopolygalacturonase I and II, J. Biol. Chem. 276:33652-33656 (2001).
66. J.A. Campbell, G.J. Davies, V. Bulone, and B. Henrissat, A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities, Biochem. J. 326:929-942 (1997).
67. B. Henrissat and G.J. Davies, Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. Plant Physiol 124:1515-1519 (2000).

68. D.T. Jones, GenTHREADER: An Efficient and Reliable Protein Fold Recognition Method for Genomic Sequences, J. Mol. Biol. 287:797-815 (1999).
69. L.J. McGuffin, K. Bryson, and D.T. Jones, The PSIPRED protein structure prediction server Bioinform. Applic. Note. 16:404-405 (2000).
70. K. Persson, H.D. Ly, M. Dieckelmann, W.W. Wakarchuk, S.G. Withers, and N.C.J. Strynadka, Crystal structure of the retaining galactosyltransferase LgtC from *Neisseria meningitidis* in complex with donor and acceptor sugar analogs, Nature Structural Biology 8:166-175 (2001).
71. N.J. Price, W.-D. Reiter, and N.V. Raikhel, Plant glycosyltransferases, Curr. Opin. Plant Biol. 2001:219-224 (2001).
72. S.J. Charnock, B. Henrissat, and G.J. Davies, Three-dimensional structures of UDP-sugar glycosyltransferases illuminate the biosynthesis of plant polysaccharides, Plant Physiol. 125:527-531 (2001).
73. U.M. Ünligil and J.M. Rini, Glycosyltransferase structure and mechanism, Current Opinion in Structural Biology 10:510-517 (2000).
74. E. Boix, Y. Zhang, G.J. Swaminathan, K. Brew, and K.R. Acharya, Structural basis of ordered binding of donor and acceptor substrates to the retaining glycosyltransferase, • -1,3-galactosyltransferase, J. Biol. Chem. 277:28310-28318 (2002).
75. C.M. Cabral, Y. Liu, K.W. Moremen, and R.N. Sifers, Organizational diversity among distinct glycoprotein endoplasmic reticulum-associated degradation programs, Mol. Biol. Cell 13:2639-2650 (2002).
76. M. Mallisard, S. Zeng, and E.G. Berger, Expression of functional soluble forms of human β 1,4-galactosyltransferase I, α-2,6-sialyltransferase, and α-2,6-sialyltransferase, and α-1,3-fucosyltransferase VI in the methylotrophic yeast *Pichia pastoris* , Biochem. Biophys. Re Commun. 267:169-173 (2000).
77. P.A. Romero, M. Lussier, A.M. Sdicu, H. Bussey, and A. Herscovics, KTR1P is an alpha-1,2-mannosyltransferase of *Saccharomyces cerevisiae* comparison of the enzymatic properties of soluble recombinant KTR1P and KRE2P/MNT1P produced in *Pichia pastoris*, Biochemical Journal 321:289-295 (1997).
78. U. Hochstrasser, M. Luscher, C. DeVirgilio, T. Boller, and A. Wiemken, Expression of a functional barley sucrose-fructan 6-fructosyltransferase in the methylotrophic yeast *Pichi pastoris*, Febs Letters 440:356-360 (1998).
79. P.F. Gallet, H. Vaujour, J.-M. Petit, A. Maftah, A. Oulmouden, R. Oriol, C. Le Narvor, M. Guilloton, and R. Julien, Heterologous expression of an engineered truncated form of human Lewis fucosyltransferase (Fuc-TIII) by the methylotrophic yeast *Pichia pastoris*, Glycobiology 8:919-925 (1998).
80. K.W. Moreman and O. Touster, Topology of mannosidasw II in rat liver membrane and release of the caalytic domain by selective proteolysis, Jr Biol. Chem 23:10945-10951 (1986).
81. R.K. Merkle, Y. Zhang, P.J. Ruest, Y.F. Liao, and K.W. Moremen, Cloning, expression, purification, and charactarization of the murine lysosomal acid alpha-mannosidase, Biochim. Biophys. Acta 1336:132-146 (1997).
82. Y.F. Liao, A. Lal, and K.W. Moremen, Cloning, expression, purification, and characterization of the human broad-specificity lysosomal acid alpha-mannosidase, J. Bi Chem. 271:28348-28358 (1996).

83. M. Bar-Peled, E. Lewinsohn, R. Fluhr, and J. Gressel, UDP-rhamnose:flavanone-7-*O*-glucoside-2"-*O*-rhamnosyltransferase. Purification and characterization of an enzyme catalyzing the production of bitter compounds in citrus, J. Biol. Chem. 266:20953-20959 (1991).
84. S.S. Basu, G.D. Dotson, and C.R.H. Raetz, A facile enzymatic synthesis of uridine diphospho-[14C]galacturonic acid, Anal. Biochem. 280:173-177 (2000).
85. H.F. Quigley, D.L. Howard, J. Sterling, and D. Mohnen, A non-radioactive gel electrophoresis assay for homogalacturonan α-1,4-galacturonosyltransferase (in preparation)
86. M. Bao, J.L. Booth, B.J. Elmendorf, and W.M. Canfield, Bovine UDP-*N*-acetylglucosamine:lysosomal-enzyme *N*-acetylglucosamine-1-phosphotransferase. I. Purification and subunit structure, J. Biol. Chem. 271:31437-31445 (1996).
87. J.P. Briand, C. Barin, M.H. van Regenmortel, and S. Muller, Application and limitations of the multiple antigen peptide (MAP) system in the production and evaluation of anti-peptide and anti-protein antibodies, J. Immunol. Methods 156:255-265 (1992).
88. V. Ziehl, R.P. Clay, C. Bergmann, A. Darvill, and P. Albersheim, Polyclonal antibodies directed against synthetic N-terminus of fungal endopolygalacturonase recognizes the native protein and cross-reacts with endopolygalacturonase of a different funal species (in preparation)
89. E. Harlow and D. Lane, Antibodies: A laboratory manual, Cold Spring Harbor Press, Cold Spring Harbor, NY, (1988).
90. D. Mohnen, H. Shinshi, G. Felix, and F.J. Meins, Hormonal regulation of β-1,3-glucanase messenger RNA levels in cultured tobacco tissues. EMBO J. 4:1631-1635 (1985).
91. G.F. Zhang and L.A. Staehelin, Functional compartmentation of the golgi apparatus of plant cells; Immunocytochemical analysis of high-pressure frozen- and freeze-substituted sycamore maple suspension culture cells, Plant Physiol. 99:1070-1083 (1992).
92. K.J. Colley, Golgi localization of glycosyltransferases: more questions than answers, Glycobiology 7:1-13 (1997).
93. A.S. Opat, C. van Vliet, and P.A. Gleeson, Trafficking and localization of resident Golgi glycosylation enzymes, Biochemie 83:763-773 (2001).
94. D. Dirnberger, P. Bencur, L. Mach, and H. Steinkellner, The Golgi localization of Arabidopsis thaliana β1,2-xylosyltransferase in plant cells is dependent on its cytoplasmic and transmembrane sequences, Plant Mol. Biol. 50:273-281 (2002).
95. S. Munro, An investigation of the role of transmembrane domains in Golgi protein retention, The EMBO Journal 14:4695-4704 (1995).
96. T. Nilsson, P. Slusarewicz, M.H. Hoe, and G. Warren, Kin recognition. A model for the retention of Golgi enzymes. FEBS Lett. 330:1-4 (1993).
97. G. Freshour, R.P. Clay, M.S. Fuller, P. Albersheim, A.G. Darvill, and M.G. Hahn, Developmental and tissue-specific structural alterations of the cell-wall polysaccharides of *Arabidopsis thaliana* Roots, Plant Physiol. 110:1413-1429 (1996).
98. A.L. Samuels, T.H. Giddings, and L.A. Staehelin, Cytokinesis in tobacco BY-2 and root tip cells: a new model of cell plate formation in higher plants, J. Cell Biol. 130:(1995).
99. B.C.-H. Lam, T.L. Sage, F. Bianchi, and E. Blumwald, Role of SH3 Domain-Containing Proteins in Clathrin-Mediated Vesicle Trafficking in Arabidopsis, Plant Cell 13:2499-2512 (2001).

100. C.-F. Chuang and E.M. Meyerowitz, Specific and heritable genetic interference by double-stranded RNA in arabidopsis thaliana, Proc. Natl. Acad. Sci. USA 97:4985-4990 (2000).
101. A.F. Bent, B.N. Kunkel, D. Dahlbeck, K.L. Brown, R. Schmidt, J. Giraudat, J. Leung, and B.J. Staskawicz, *RPS2* of *Arabidopsis thaliana*: A Leucine-Rich Repeat Class of Plant Disease Resistance Genes, Science 265:1856-1860 (1994).
102. E. Bieberich, S. MacKinnon, J. Silva, D.D. Li, T. Tencomnao, L. Irwin, D. Kapitonov, and R.K. Yu, Regulation of Ganglioside Biosynthesis by Enzyme Complex Formation of Glycosyltransferases, Biochem. 41:11479-11487 (2002).
103. B.L. Ridley, M.D. Spiro, J. Glushka, P. Albersheim, A. Darvill, and D. Mohnen, A method for biotin labeling of biologically active oligogalacturonides using a chemically stable hydrazide linkage, Anal. Biochem. 249:10-19 (1997).
104. F. Guillaumie, K.J. Jensen, O.R.T. Thomas, J.D. Sterling, and D. Mohnen, Solid-phase biosynthesis and MALDI-TOF mass spectrometry analysis of pectic oligogalacturonides: a new tool to monitor the extension of a homogalacturonan chain (in preparation)

PART II

Functional Identification of Genes for Pectin Biosynthetic Enzymes

Project Summary

List of Senior Personnel

Debra Mohnen, PI, CCRC and Department of Biochemistry and Molecular Biology, University of Georgia
Maor Bar-Peled, Co-PI, CCRC and Department of Plant Biology, University of Georgia
Michael Hahn, Co-PI, CCRC and Department of Plant Biology, University of Georgia
Malcolm A. O'Neill, Co-PI, CCRC, University of Georgia
William S. York, Co-PI, CCRC and Department of Biochemistry and Molecular Biology, University of Georgia

Summary of proposed project

The enzyme activity of 32 proteins encoded by putative pectin biosynthetic genes will be determined. Pectin is a family of structurally complex cell wall polysaccharides that has multiple roles in plant growth, development and disease resistance. Pectin synthesis is catalyzed by glycosyltransferases that transfer a glycosyl residue from an activated sugar onto an oligosaccharide/polysaccharide acceptor. The putative pectin biosynthetic genes were identified based on: (1) sequence similarity to the recently identified gene (*GALAT1*), an *Arabidopsis thaliana* UDP-galacturonic acid: HGA α-1,4-galacturonosyltransferase (GALAT1); (2) sequence similarity to KDOtransferases; and (3) sequence similarity to a putative glucuronosyltransferase. To identify enzyme activity (1) *each of the genes will be cloned, heterologously expressed, and the expressed protein tested in a multi-acceptor pectin biosynthesis assay (MAPA);* (2) *the visible phenotype and wall pectin structure of mutant plants with T-DNA inserts in the selected genes will be characterized;* (3) *polyclonal antibodies to heterologously-expressed pectin biosynthetic enzymes will be generated and used to define the subcellular location of the gene products and to inhibit specific pectin biosynthesis activity in permeabilized-membranes that contain endogenous acceptors.* The long-term goals are to identify and characterize genes involved in pectin biosynthesis.

Broader Impacts of the proposed research

The proposed research will lead to the functional characterization of *Arabidopsis* genes that catalyze pectin synthesis. The genes and their confirmed enzyme activities will be available at a new web link entitled "Pectin Biosynthesis NSF Arabidopsis 2010 Project" on the CCRC homepage (http://www.ccrc.uga.edu). Acceptor molecules and cDNA clones generated during this study will be made available to researchers through a "Pectin Biosynthesis Resource Manager" funded from this proposal and using the infrastructure of CarboSource Services, a component of the NSF-funded Plant Cell Wall Biosynthesis Research Network (http://xyloglucan.prl.msu.edu/). The identification of genes encoding pectin biosynthetic enzymes will allow the engineering of plants to produce pectins with modified structures and properties, pectins with improved agricultural value, and pectin-based neutraceutical and pharmacological products. The participation of historically underrepresented groups in the research will be achieved by recruiting undergraduate and graduate students through the University of Georgia Summer Undergraduate Research Program (SURP) (www.gradsch.uga.edu/rr/) and the new SURP-Bridge program, which draw in minority students from across Georgia. The latter program offers

incoming graduate students from underrepresented groups research experience prior to graduate school in an effort to develop the students' research skills and increase their likelihood of success in graduate school.

I. Specific Aims

We have identified 32 genes in *Arabidopsis* (see Appendix A-1) that we predict encode glycosyltransferases involved in pectin biosynthesis. The putative pectin biosynthetic genes to be investigated in this proposal include a family of 15 putative GalAT genes and a family of 10 GalAT-like genes identified based on their sequence similarity to the functionally and biochemically identified *GALAT1* (Sterling and Mohnen, in preparation). Two other families of genes were selected based on sequence similarity to a putative RG-II glucuronosyltransferase[1] (the GlcAT family) and to prokaryotic KDOtransferases[2-4] (the KdoT family). The goal of this proposal is to determine the catalytic activity and substrate specificity of the proteins encoded by these genes. At least 50 unique glycosyltransferases are likely to be required to synthesize pectin[5]. The number of pectin biosynthetic glycosyltransferases may be significantly larger if multiple copies of the enzymes are encoded, as is likely since at least parts of the *Arabidopsis* genome appear to have undergone duplication[6].

We recently identified the *Arabidopsis* gene (*GALAT1*) that encodes α1,4-galacturonosyltransferase I (GALAT1), a pectin biosynthetic glycosyltransferase. GALAT1 transfers galacturonic acid (GalA) from UDP-GalA onto homogalacturonan acceptors[7-10]. Thus, *GALAT1* encodes a pectin biosynthetic glycosyltransferase. We will use the methodology we developed during the identification of *GALAT1* to characterize the enzyme activity of the putative pectin biosynthetic genes described in this proposal. All 32 genes will be (1) cloned, heterologously expressed as epitope-tagged constructs to facilitate purification, and when applicable as N-terminally truncated constructs, and transformed into *Pichia* and human embryonic kidney cells (HEK293 cells). Enzyme activity of the expressed recombinant proteins will be determined using a multi-acceptor pectin biosynthesis assay (MAPA) that allows screening for multiple pectin biosynthetic activities. (2) Antibodies will be generated against each heterologously expressed protein. For proteins that give no activity in the MAPA, we will determine whether the specific antibody can inhibit pectin biosynthesis in permeabilized membrane preparations (permeabilized membranes assay, PMA) that have endogenous acceptors[8,10] rather than exogenous acceptors. Antibody-dependent inhibition of specific enzyme activity will allow us to identify the nucleotide-sugar and/or acceptor substrates for the expressed enzymes and allow us to optimize the acceptor substrates used for MAPA. The antibodies will also be used to determine the subcellular localization of the gene products using tissues shown, by RT-PCR, to express the mRNA for the specific gene. The emphasis will be on immunogold localizations using electron microscopy, since the subcellular location of the genes will be most informative regarding the possible role the gene products in pectin biosynthesis. (3) We also propose to structurally characterize the pectin in the walls of *Arabidopsis* plants with T-DNA inserts in the genes of interest. Changes in pectin structure in the mutants would indicate the defect in wall synthesis, give information about the enzyme activity of the gene product, and allow us to isolate pectic fragments from the mutants for use as additional exogenous acceptors in the MAPA. To date, T-DNA insertion mutant lines for 19 of the selected genes are available in the Salk *Arabidopsis* T-DNA library (http://signal.salk.edu/cgi-bin/tdnaexpress).

The genes to be tested in this proposal, and the confirmed enzyme activity of the gene products, will be listed at a web link entitled "Pectin Biosynthesis NSF Arabidopsis 2010 Project" to be established at the CCRC web site (http://www.ccrc.uga.edu). Acceptor substrates and cDNA clones will be made available to other researchers through a "Pectin Biosynthesis Resource Manager" to be funded from this proposal and using the infrastructure of the established CarboSource Services, a component of the NSF-funded Plant Cell Wall Biosynthesis Research Network (http://www.ccrc.uga.edu/web/services/carbosource/index.html). The combined expertise of the PI and Co-PIs in pectin structure[11-15], biosynthesis[5,7-11,16-19] and function[20-27], as well as carbohydrate chemistry,

biochemistry, molecular and cell biology[28-36] and immunochemistry [37-40] provides a unique breadth and strength in multiple research disciplines and methodologies to tackle the challenging task of identifying pectin biosynthetic genes. The identification of such genes will allow the eventual engineering of plants that produce modified pectin, that have improved agricultural value, and will likely allow the production of pectin-based neutraceutical[41,42] and pharmacological[42-45] products.

II. Significance

Pectin is the most structurally complex polysaccharide in the plant cell wall. It comprises 30-40% of the primary wall of dicots and non-graminaceous monocots, and ~10% of the primary wall in the grass family. Pectins are a family of polysaccharides[11,14,18] that include homogalacturonan (HGA) (Fig. 1), rhamnogalacturonan-I (RG-I) (Fig. 2) and the substituted galacturonans including the ubiquitous rhamnogalacturonan II (RG-II) (Fig. 3), xylogalacturonan (XGA)[46-48] and apiogalacturonan[49,50]. Pectic polysaccharides impart structure[20] to the growing plant cell wall, regulate ion binding in the wall, influence cell-cell adhesion[60,70], are involved in cell signaling[22,57], and may control the diffusion of molecules through the apoplast[51]. Pectins have been implicated in a broad range of plant growth phenomena including pollen tube growth[52], seed hydration[53,54], leaf abscission[55], water movement[56], and fruit development[11]. In addition, pectic oligosaccharides serve as signals[22,57] during plant development[22] and induce plant defense responses[58,59]. Mutant studies have shown that altered pectin structure leads to dwarfed plants[20], brittle leaves[60], reduced numbers of side shoots and flowers[61], and plants with reduced cell-cell adhesion[1,62]. The successful completion of the proposed research will lead to the identification of the enzyme activity, and thus biochemical function, of multiple pectin biosynthetic genes, many of which are likely to be involved in the synthesis of the backbones of the three pectic polysaccharides HGA, RG-I and RG-II. The availability of these genes would provide the molecular and biochemical tools needed to identify additional glycosyltransferases involved in branching of the backbones, and would allow the generation of plants with altered pectin structure. While the 32 genes to be studied in this proposal represent only ~0.12% of the ~28,000 genes in *Arabidopsis*, they are some of the most difficult genes to identify and characterize because of a lack of commercially available acceptor substrates and activated glycosyl donor substrates.

The *GALAT1* gene has high sequence similarity to proteins expressed in other plants, thus it is likely that the other pectin biosynthetic genes will also have identifiable homologs in other plant species, including agriculturally important plants. Since pectin of overall very similar structure is present in the walls of all flowering plants and gymnosperms, the identification of functional pectin biosynthetic genes will greatly facilitate the engineering of plants with modified pectin and with altered growth characteristics, some of which are expected to yield plants of increased agronomical value. In addition, mutant plants with defined changes in pectin synthesis will allow the dissection of the biological role of each pectic component in plants. The pectin biosynthetic genes will also provide valuable tools for understanding mechanistically how pectin is synthesized. The glycosyltransferase-specific antibodies to be generated will also allow the process of pectin assembly in the Golgi to be elucidated. A complete understanding of such a polysaccharide cellular trafficking process is unknown in any biological system.

None of the 32 selected putative pectin biosynthetic genes are listed among the currently funded 2010 projects except for the large scale 2010 projects by Ecker ("A Sequence-Indexed Library of Insertion Mutations in the Arabidopsis Genome Sequence-Indexed TDNA Library", NSF#0115103) and Ecker/Davis/Theologis ("Global Expression Studies of the Arabidopsis Genome, NSF#0196098). At1g18580 is listed in the project entitled "Expression Profiling of Plant Disease Resistance Pathways" (NSF#0114783) by Dong, a project that includes a list of ~2000 genes. It is unlikely that the Dong study will test enzyme activity of At1g18580 and thus, our proposed research compliments that of Dong.

III. Relation to longer-term goals of the PI and Co-PIs

The PI and Co-PI's are faculty at the Complex Carbohydrate Research Center and all have a long track-record in studying pectin structure[11-15], function[20-27] and biosynthesis[5,7-11,16-19]. The long-term goals

of the combined group is to understand the structure of the complex family of pectic polysaccharides, to decipher how wall polymers, including pectin, are synthesized and assembled into a growing wall, and to determine the function of pectin in plants using physical/chemical, biochemical, genetic and molecular and cell biology approaches. Thus, for the PI and Co-PI's the current proposal represents a gene discovery effort to supply tools (i.e. genes for biosynthetic enzymes) that will allow detailed mechanistic studies on the biology of plant cell walls.

IV. Introduction

IV.A. Pectin Structure

Pectin is a family of structurally complex polysaccharides that contain 1,4-linked α-D-galacturonic acid residues[11,14,18]. Together they account for 30-40% of the primary wall of dicots and non-graminaceous monocots, and ~ 10% of the primary wall in the grass family. Pectin is comprised of homogalacturonan (HGA) (Fig. 1), rhamnogalacturonan I (RG-I) (Fig/ 2) and the substituted galacturonans rhamnogalacturonan II (Fig. 3), xylogalacturonan (XGA)[46-48] and apiogalacturonan4[9,50] (see pdf file entitled "Pectin Review" at http://www.ccrc.uga.edu/~dmohnen/nsf2010pdf/reviewers.html) for summary of pectin structure[11]). HGA accounts for 55%-70% of pectin[47]. HGA is a linear homopolymer of 1,4-linked α-D-galactosyluronic acid that is often partially methylesterified at the C6 carboxyl group and may be partially acetylated at O-2 and/or O-3[11] (Fig. 1). RG-II is a structurally complex polysaccharide that accounts for approximately 10% of pectin[11,63]. RG-II has an HGA backbone that is substituted at O-2 and/or 0-3 with four structurally complex oligosaccharides[11,14] (Fig. 3). RG-II contains 11 different types of sugars including D-glucuronic acid (GlcA), D-Kdo (ketodeoxymanno-octulopyranosylonic acid), and D-Dha (deoxylyxo-heptulopyranosylaric acid). Rhamnogalacturonan I (RG-I) accounts for 20-35% of pectin[63] (Fig. 3). RG-I is a family of polysaccharides with an alternating [→4)-α-D-GalA-(1→2)- α-L-Rha-(1→] backbone in which between 20-80% of the rhamnoses are substituted predominantly by arabinan, galactan, or arabinogalactan side branches[11,18,64]. Some of the side chains are terminated with GlcA or 4-0-methyl-GlcA residues[65,66]. The composition and length of the side chains in RG-I varies between cell types and in different plant species[11,67]. The walls of some plants (e.g. *Lemna*) contain apiogalacturonans (AGA) in which the galacturonan backbone is substituted at the 2 or 3 position by D-apiofuranose[49,50]. Xylogalacturonans (XGA) have a backbone that is substituted at 0-3 with D-xylose and are often found in reproductive tissues[47,48,68,69].

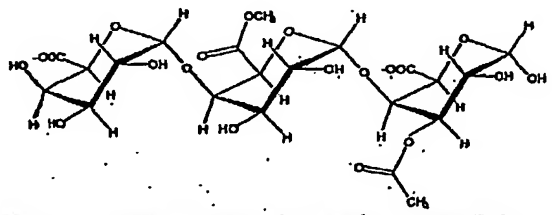


Figure 1. ***Trimeric region of homogalacturonan (HGA).*** HGA is a linear homopolymer of α-1,4-linked galacturonic acid that may be methylesterfied at C6 and acetylated at O2 or O3. Substituted galacturonans, such as RG-II, apiogalacturonan and xylogalacturonan have an HGA backbone.

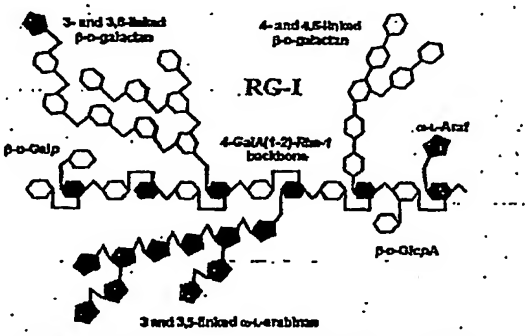


Figure 2. ***Represenative structure of rhamnogalacturonan I (RG-I).*** RG-I has an alternating [→4)-α-D-GalpA-(1→2)- α-L-Rhap-(1→] backbone in which roughly 20-80% of the rhamnoses are substituted by arabinans, galactans, or arabinogalactans.

Although pectins appear to have multiple roles in plants, the specific mechanism of action has not been determined. To determine the biological roles of pectins requires the generation of plants with specific alterations in pectin structure by knocking out genes that encode pectin biosynthetic enzymes.

Such enzymes include the nucleotide-sugar biosynthetic enzymes and the pectin glycosyltransferases. Each glycosyltransferase is expected to transfer a unique glycosyl residue in a specific linkage in an anomer and linkage-specific manner onto a specific polymeric/oligomeric acceptor. Immunohistochemical and biochemical studies indicate that pectin synthesis occurs in the Golgi apparatus[71-74]. Pectin biosynthetic genes are most likely sequentially expressed in different Golgi cisternae[75,76] in the order in which pectin is synthesized as it moves from the cis, through the medial and to the trans Golgi.

To date, only five[31,77-79] of the more than 200 predicted wall biosynthetic glycosyltransferases have been functionally identified at the gene level (i.e. enzyme activity of the encoded protein confirmed) and none of these encode pectin biosynthetic genes. We have recently used a partial purification-tandem mass spectrometry approach to identify the first gene (*GALAT1*) encoding a demonstrated pectin biosynthetic enzyme, GALAT1 (see below). The identification of this gene allowed us to identify a putative GalAT gene family and a putative GalAT-like gene family in *Arabidopsis*.

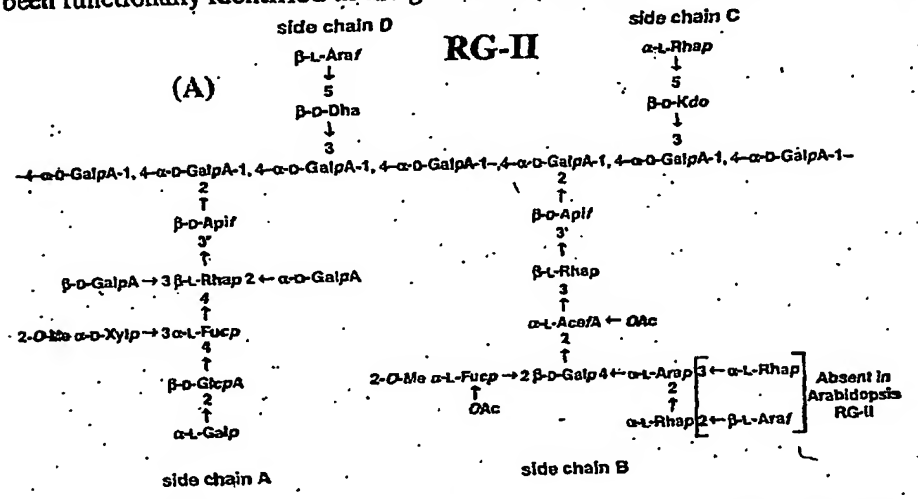


Figure 3. ***Representative structure for rhamnogalacturonan II*** **(RG-II).** RG-II has a backbone of 1,4-linked α-D-Gal*p*A residues. GalA residues are also present in RG-II side chain B.

IV. B. Pectin Synthesis

At least 14 distinct enzyme activities are required to synthesize the nucleotide-sugar substrates and 58 distinct glycosyl-, methyl- and acetyl-transferases are required to synthesize pectin[5] (for pdf file of review see "Pectin Biosynthesis" at http://www.ccrc.uga.edu/~dmohnen/nsf2010pdf/reviewers.html). Between 4-9 unique galacturonosyltranferases are predicted to be required for the synthesis of HGA, RG I, RG-II and possibly for the synthesis of XGA and AGA[5]. Since, the *GALTA1* gene was identified based on the presence of the GALAT1 protein in partially purified solubilized *Arabidopsis* membrane protein fractions that contained UDP-GalA:HGA α1,4GalAT enzyme activity, we propose that GALAT1 is UDP-GalA:HGA α1,4GalAT.

Membrane-bound α1-4galacturonosyltransferase (GalAT) activity has been identified and partially characterized in numerous plants [7,10,80-85] including *Arabidopsis* (Sterling and Mohnen, unpublished results). Pea GalAT was localized to the Golgi with its catalytic site facing the lumenal side of the Golgi providing the first direct enzymatic evidence that the synthesis of HGA occurs in the Golgi[10]. *In vitro*, GalAT adds [^{14}C]GalA from UDP-[^{14}C]GalA[19,86] onto endogenous acceptors in microsomal membrane preparations to produce radiolabeled products with molecular masses between 100 and 500 kd[7,10]. Treating the radiolabeled product with a purified endopolygalacturonase generated GalA, digalacturonic acid and trigalacturonic acid, thereby confirming that the synthesized product was HGA. Thus, *in vitro* the enzyme catalyzes the reaction: UDP-GalA + HGA(n) → HGA(n+1) + UDP, where HGA can be pure HGA or HGA regions in pectin. The product produced *in vitro* in tobacco microsomes was ~ 50% esterified[7] whereas the product produced in pea Golgi did not appear to be heavily esterified[10], suggesti

that the degree of methyl esterification of newly synthesized HGA may be species specific and that methylesterification occurs after synthesis of at least a short stretch of HGA.

Enzymatically active GalAT is solubilized from membranes with detergent[8]. Solubilized GalAT adds GalA onto the non-reducing end[9] of exogenous HGA (oligogalacturonide; OGA) acceptors of a degrees of polymerization $\geq 10^7$. Solubilized GalAT from tobacco membranes[8], or detergent-permeabilized Golgi from pea[10], elongates exogenous OGAs predominantly by a single GalA residue unless high UDP-GalA:OGA ratios are used[8,11,84,87], suggesting that solubilized GalAT *in vitro* acts nonprocessively, (i.e. distributively). The apparent lack of *in vitro* processivity of GalAT was recently confirmed by Akita *et al.*[85] who showed that a solubilized enzyme from petunia pollen can elongate OGAs in a "successive" fashion with up to 10 GalA residues, but the kinetics of the reaction suggested a distributive (i.e. non-processive) mode of action..

We identified a putative GalAT gene from *Arabidopsis* using a partially purified enzyme together with tandem mass spectrometry. *Arabidopsis* GalAT was partially purified from a detergent-solubilized enzyme preparation by sequential chromatography over ion exchange and affinity resins. Fractions enriched in GalAT activity were treated with trypsin and the amino acid sequences of the peptides determined by liquid chromatography-tandem mass spectrometry. These amino acid sequences were then used to screen the *Arabidopsis* gene/protein database. Thirty unique proteins were solely identified in the GalAT-containing fractions. Among the 30 unique proteins that co-purified with GalAT activity, two proteins (designated JS33 and JS36) were identified as putative GalAT proteins/genes based on their having at least one predicted transmembrane domain and a predicted glycosyltransferase domain (see CAZy database; http://afmb.cnrs-mrs.fr/CAZY/index.html).

JS33 and JS36 were either cloned by RT-PCR (JS36) using total RNA from *Arabidopsis* tissue or a cDNA clone was obtained (JS33) from the Arabidopsis Biological Resource Center. The proteins encoded by these genes each have a predicted single transmembrane domain and a high pI as expected based on the biochemical characteristics of the GalAT enzyme[5,7,8]. The genes were truncated to remove their N-terminal region, including all or most of the predicted transmembrane domain and the truncated genes were inserted into the mammalian expression vector pEAK, that contained a heterologous signal sequence, a histidine (HIS) tag, and two influenza hemagglutenin (HA) epitopes.

The truncated forms of JS33 and JS36 in the pEAK vector were transiently expressed for 46 hours in human embryonic kidney cells (HEK293 cells). The media was collected and incubated with a mouse anti-HA IgG that was bound to Protein A Sepharose since the translational fusion protein constructs contained two copies of the HA epitope N-terminal to the predicted catalytic domain. The immunoabsorbed protein was assayed for GalAT activity. Table I shows that the JS36 construct expressed GalAT activity. These results show that JS36 is a GalAT and we have named the gene encoding this enzyme *GALAT1* (Sterling *et al.*, in preparation) in accordance with the rules put forward by the "Commission on Plant Gene Nomenclature (http://genome-www.stanford.edu/Mendel/) and The Arabidopsis Information Resource (TAIR) (http://www.arabidopsis.org/nomencl.html; http://www.arabidopsis.org/info/guidelines.html#genes). Stable transformed lines of JS36 and JS33 are currently being established to produce large amounts of heterologously expressed GALAT1 for further study.

Analysis of the gene sequence of GALAT1 shows that the predicted expressed protein contains a transmembrane domain. This is in agreement with the GalAT activity being membrane bound in all species examined to date[5]. Furthermore, the predicted topology of GALAT1 is that of a type-II membrane protein, in agreement with our previous studies showing that the catalytic site of pea GalAT is in the lumen of the Golgi[10]. Type-II glycosyltransferase membrane proteins have a short N-terminal cytosolic tail, a transmembrane region, a stem region, and a C-terminal catalytic domain[89].

GALAT1 is a member of Glycosyltransferase Family 8 of the CAZy database [a database of putative and proven carbohydrate modifying enzymes that currently has 61 different glycosyltransferase families (http://afmb.cnrs-mrs.fr/CAZY/index.html)[90,91]]. The presence of GALAT1 in Family 8 is

consistent with our demonstrated activity of GALAT1 as an α1,4-galacturonosyltransferase, since Family 8 members are retaining glycosyltransferases. GALAT1 is a retaining enzyme because the α-configuration in the substrate UDP-α-GalA is retained in the product α1,4-linked-galacturonoman (HGA).

Table I. Expression of recombinant putative GalAT in mammalian HEK293 cells. HA-epitoped tagged truncated JS36 and JS33 genes inserted into the mammalian expression vector pEAK, along with a pEAK vector only control, were independently transiently transfected into HEK293 cells. Transfected cells were grown for 46 hours and the media was incubated with anti-HA IgG bound to either Protein A Sepharose. Immunoabsorbed protein was assayed for GalAT activity using UDP-[^{14}C]GalA and oligogalacturonides as substrates. The data are the average [^{14}C]GalA incorporated into product (minus T0 controls) from duplicate 45 min reactions. Time 0 controls: reactions to which base was added to inactivate the enzyme prior to the addition of substrates. Positive control: GalAT activity in solubilized *Arabidopsis* protein. Similar results were obtained in two independent experiments.

Enzyme	[^{14}C]GalA incorporated
Recombinant truncated JS36	161
Recombinant truncated JS33	30
Vector control	0
Solubilized Arabidopsis enzyme (control)	1434

GALAT1 is expressed in all *Arabidopsis* tissues (flower, root, stem and leaf) we examined by RT-PCR (Sterling *et al.*, in preparation). Furthermore, 18 EST entries for this gene in the TAIR database (http://www.arabidopsis.org/) indicate that GALAT1 mRNA is expressed in *Arabidopsis* developing seed, green siliques, roots and aerial organs.

V. Proposed research

V. A. Selection of putative pectin biosynthetic genes

V.A.1. Identification of a proposed GalAT Gene Family. We propose that GALAT1 and the 14 other *Arabidopsis* genes with high sequence similarity to GALTA are members of a *GalAT family* and are involved in HGA, RG-I, RG-II and XGA synthesis.

A standard protein blast and a PSI Blast of the NCBI protein database using the GALAT1 (JS36) amino acid sequence revealed that GALAT1 is one member of a 15 member putative GalAT gene family in *Arabidopsis*. The genes we selected for this GalAT family share ≥ 30% amino acid identity and ≥ 50% amino acid similarity based on a PSI Blast. These genes were further compared with GALAT1using a Pairwise BLAST. The results indicate that this family of genes has ≥ 34% identity and ≥ 52% similarity to GALAT1 in the portion of the genes C-terminal to the membrane spanning domain. This identity is comparable to the 37-54% identity shared among the proposed ten member *Arabidopsis* fucosyltransferase gene family (AtFUT1-10)[92].

GALAT1 is the only member of the GalAT family that has a known function. However, the recent characterization of the mutant *Quasimodo1*[62], is in agreement with our identification of GALAT1 as a galacturonosyltransferase. The GalA content of the walls of *Quasimodo1* is reduced by 25% and plants exhibit decreased cell adhesion[62], characteristics consistent with the possibility that the *Quasimodo1* gene encodes a GalAT. *Quasimodo1* has 53% amino acid identity and 72% similarity to GALAT1 and the gene affected in *Quasimodo1* is a member of our proposed putative GalAT family. There is as yet, however, no direct evidence that the protein encoded by *Quasimodo1* is a functional GalAT.

The conserved amino acids within the GalAT gene family are shown in Fig. 4. Glycosyltransferases typically contain one or more carboxylates at the catalytic site. At least one of these carboxylates is expected to coordinate a divalent cation associated with the nucleotide-sugar. In many glycosyltransferases the metal coordination involves two carboxylates that are often present as

DDx, xDD, or DDD (the so-called "D(x) D" motif)[93]. We predict that D498 and D500 are part of the "D(x) D' motif. However, structural and mutational data is required to confirm this.

GALAT1 homologs are present in other plants, as would be expected for HGA GalAT. These include proteins in *Cicer arietinum* (chickpea) (76% amino acid identity; 87% similarity), *Oryza sativa* (japonica rice) (59% identify; 75% similarity) and *Populus alba* (49% identity; 72% similarity). In addition, tentative contigs encoding apparent orthologues to GALAT1, derived from EST sequencing efforts, have been identified in a number of plants including soybean, *Medicago truncatula*, and cotton.

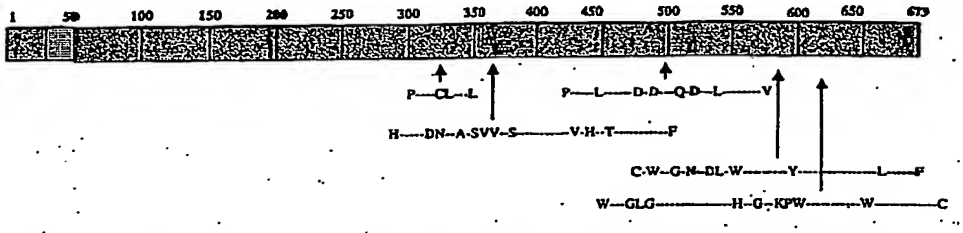
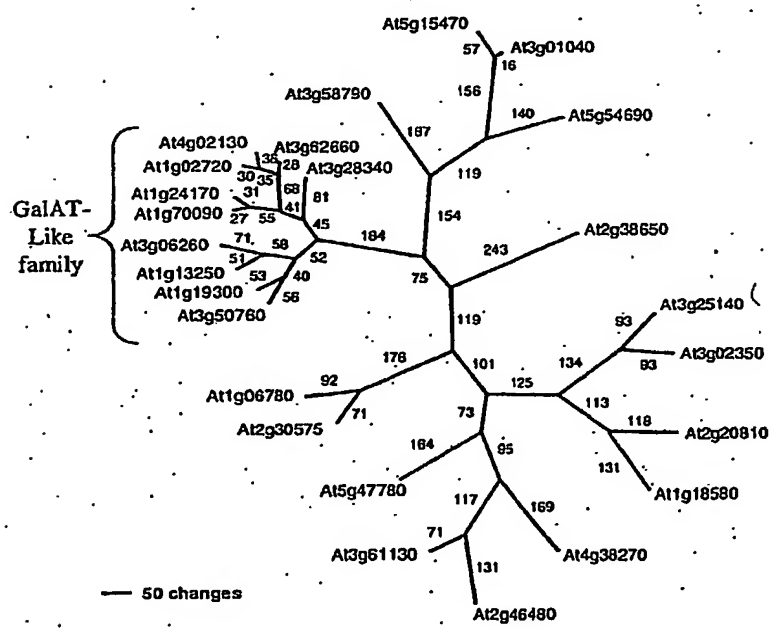


Figure 4. *Schematic representation of the conserved sequences in the Arabidopsis GalAT gene family.* The speckled region represents the transmembrane spanning region. The relative position of conserved residues represents numbering as for *GALAT1* (At3g361130).

V.A.2. GalAT-like family. A PSI Blast against GALAT1 identifies a further 10 genes that also have high sequence identity (23-29%) and similarity (41-51)% to GALAT1 and form a tight cluster of highly similar genes (55-66% identity/67-77% similarity). A Neighbor Joining Tree of our proposed *Arabidopsis* GalAT Superfamily (i.e. the proposed GalAT family and the GalAT-Like family), based on a sequence alignment generated by ClustalX[128], is shown in Fig. 5. The 10 GalAT-like genes are all significantly smaller, lacking ~200 amino acids in comparison with the GalAT family. Nonetheless, they appear to be targeted to the secretory pathway based on annotation of the genes at the Arabidopsis Information Resources. All 10 genes appear to be expressed in *Arabidopsis*, since they are represented by one or more ESTs in the *Arabidopsis* EST collection. The GalAT-like genes also contain some of the same conserved residues as the GalAT family, namely D-D——D—L (the predicted "D(x) D" motif) and L——-F———W—GLG———— H—G—KPW. We group the 10 GalAT-like genes into a family and propose that they encode GalATs directly involved in pectin synthesis or GalATs with, as yet, unidentified glycosylating function.

***Figure 5. Arabidopsis* GalAT Superfamily**



adhesion and reduction in GlcA content in the wall and in RG-II. The tobacco gene has high sequence similarity to a known glucuronosyltransferase in vertebrates. No enzyme activity for the expressed protein was reported, although chemical analysis suggested that the gene may encode a GlcAT involved in RG-II synthesis. We propose to test whether the apparent *Arabidopsis* homolog, At5g61840, encodes a functional RG-II:GlcAT. Two *Arabidopsis* genes have high sequence identity (93%) and similarity (95%) to *NpGUT1* and two other *Arabidopsis* genes have 42% identity and 62% similarity to *NpGUT1*. We propose that these four genes utilize UDP-GlcA as nucleotide sugar and transfer GlcA specifically to RG-II or RG-I. We group the genes as the putative pectin GlcAT family. This family shares the following three conserved sequence modules: S-VRT-NPEEADWFY-PVYTTCD; PYWNRT-G-DHFF--HDFGACFHY-EEKAI-G; R-IFCLCPLGWAPWSPRLVE-VIFGCIPVIIADDI-LPFAD.

V.A.4. Putative Kdotransferase family. Kdo and Dha are present in plant RG-II, some algal walls, and bacterial lipo-and exo-polysaccharides[94]. The genes that encode bacterial Kdotransferases have conserved sequences[4,95] and use CMP-Kdo as the activitated sugar donor. We identified homologous sequences in *Arabidopsis* and rice. These genes are members of CAZy Glycosyltransferase Family 30. Using RT-PCR we cloned and sequenced a putative *Arabidopsis* KdoT gene (Hogans, Harper, and Bar-Peled unpublished results) that encodes a putative membrane bound protein. The *Arabidopsis* gene shares 33% amino acid sequence identity and 49% similarity to the functional KdoT from prokaryotes[4].

Vertebrate sialyltransferases (CAZy Family 29) are a group of glycosyl transferases that utilize CMP-activated sugars. Oligosaccharides containing sialic acids are not present in plants. Nevertheless, sialic acid is structurally similar to Kdo and Dha. A 2,6- and 2,3-sialyltransferase from rat liver[96] and porcine submaxillary gland[97] have a common 55 amino acid sequence ("sialyl motif"[98] [CAVVGNSGTLLNSQY GDLIDKHEIVIRLNNAKTE-RFEKKVGSKTNISFINS]) presumably located in their active domain. *Arabidopsis* contains two genes that are grouped in the same glycosyltransferase family as the sialyltransferases (CAZy 29). The two *Arabidopsis* genes encode putative membrane bound proteins and have the same 55 amino acid domain present in mammalian sialyltransferases. Thus, we propose to determine if the two *Arabidopsis* genes in CAZy Family 29 are putative Kdotransferases.

V.B. Summary of scheme to identify gene function

We propose to take three complementary approaches to define the function of the putative pectin biosynthetic genes (Fig. 6). (1) Each gene will be heterologously expressed and the expressed proteins assayed for their ability to transfer radiolabeled sugar (from a nucleotide-sugar) to a series of oligo- and polysaccharide acceptors using a multi-acceptor pectin biosynthesis asaay (MAPA). (2) Antibodies will be generated to the expressed protein and used for subcellular immunocytochemistry to provide information on whether or not the location of the gene product supports a role in pectin synthesis. These antibodies will be also be used in a permeabilized membrane assay (PMA) to inhibit transfer of radiolabeled sugar to endogenous acceptors. (3) The pectins in the cell walls of available mutants for each gene will be structurally characterized. A change in pectin structure would suggest the point at which pectin synthesis was effected and thereby provide information on enzyme activity of

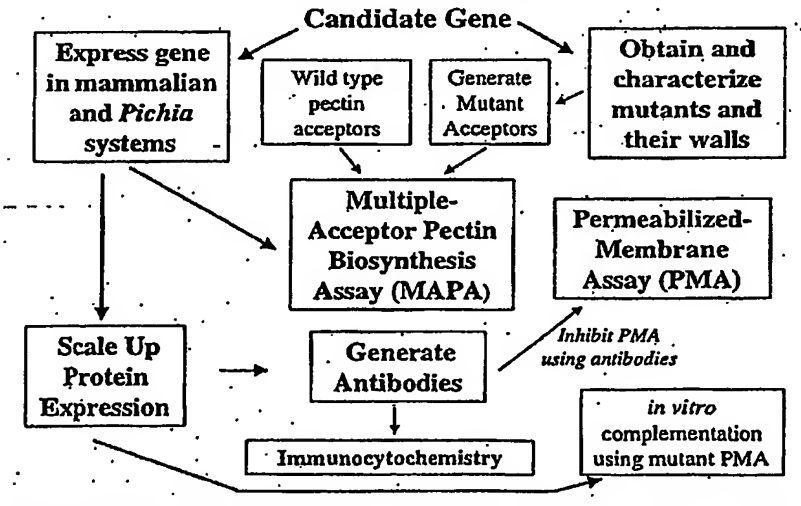


Figure 6. Summary of scheme to identify pectin biosynthetic genes.

the gene product. Pectin acceptors from such mutants would be isolated and used in the MAPA if enzymatic activity had not already been confirmed in (1). If enyzme activity for the expressed protein was not obtained, even using mutant acceptors in the MAPA, a permeabilized membrane assay (PMA) using membranes from the mutant plant and an aliquot of the heterologously expressed protein would be undertaken to attempt to add radiolabeled sugar onto the mutant acceptor and thereby complement the pectin synthesis defect in the mutant membranes. The specific methods to be used are described below.

V.C. Cloning of the genes

The selected genes will be cloned by RT-PCR using RNA from *Arabidopsis* root, stem, leaves, flower and suspension cultures (depending on where highest specific mRNA levels are expressed), or the cloned genes will be obtained from the *Arabidopsis* Biological Resource Center (ABRC) if they are already available. The genes will be inserted into the vector pCR2.1 to allow easy manipulation into other expression vectors and to facilitate DNA sequence determination. Standard procedures including PCR and restriction digestion comparisons will be used to confirm gene identify. Several of the PIs have extensive experience with these techniques [28,31,32,36,99,100].

V.D. Heterologous expression

We have shown (see section IV.B.) that human embryonic kidney (HEK293) cells transiently infected with recombinant truncated *GALAT1* expressed active recombinant GALAT enzyme. Whereas transient expression systems are sufficient to measure activity of an expressed protein, alternative expression strategies are required to produce proteins in quantities sufficient for further characterization of the enzyme and for antibody production (~100 µg/immunization). One strategy will be to produce stable transfected clonal HEK293 lines[101] expressing each of the selected genes. Truncated versions of the genes will be made by removing the N-terminus and transmembrane domain as was down for GALAT1. The gene constructs will contain an N-terminal epitope-tag since such GALAT1 constructs expressed in mammalian cells were active, and since numerous type II membrane glycosyltransferases and glycosidases have been shown to retain activity with N-terminal tags[102,103] and N-terminal truncations[31,77]. An alternative strategy will be to express the full length and truncated forms of the genes in *Pichia pastoris*. Although both *Pichia* and mammalian cells can be used to express plant wall biosynthetic enzymes[31,77,78], it is not possible to predict, *a priori*, which of the systems will produce the highest amounts of functional enzyme. Thus, we will initially make constructs in both systems.

Kelley Moremen, a faculty member at the CCRC routinely expresses proteins in mammalian cells and has advised us on the transient expression of GALAT1 in mammalian HEK293 cells. He is also advising us on our current production and cloning of stable GALAT1 transfected cell lines, and will continue to advise us for this proposal (see letter of support). The basic protocol will be as in the Edge BioSystems protocols with the modification suggested by the Moremen lab.

For expression in *P. pastoris*, cDNA encoding the entire, and the truncated soluble forms, of putative pectin biosynthetic enzymes will be generated by PCR using gene/vector specific primers. The PCR products will be subcloned into an appropriate *Pichia* expression vector in which the cDNA is inserted downstream from an alcohol oxidase (AOX1) promoter (Invitrogen). We are preparing *GALAT1* constructs in the *Pichia* and will determine whether functional C-terminal epitope-tagged constructs that do not effect GalAT activity can be recovered. If activity is recovered, we will make similar constructs for the other putative pectin biosynthetic genes. *Pichia* will be transformed with the linearized recombinant constructs and expression of the recombinant protein will be induced by growth in methanol. Cells will be plated and screened for growth under selection and the selected colonies will be assayed for enzyme activity in the MAPA. Approximately 10-20 colonies enzyme expressing clonal lines will be quantitatively assayed for their respective enzyme activity to identify the highest expressing line. Numerous glycosyltransferases[104-107] and plant enzymes [108-110] have been functionally expressed in *Pichia*, and one of the Co-PIs (Bar-Peled) has functionally expressed nucleotide-sugar biosynthetic enzymes (UDP-Xyl synthase, UDP-Rha epimerase/reductase, UDP-GlcA epimerase) in *Pichia*.

Expression in *Pichia* is more cost-effective than expression in mammalian cells and can generate high levels of expressed protein (≥1 gm L^{-1}). Once a high-enzyme expressing line is recovered, large amounts of protein will be produced in fermentors (available at the CCRC and at the University of Georgia Fermentation Research Facility) or spinner flasks and related equipment used for scale up of mammalian cell culture (available at the CCRC).

V.E. Multi-acceptor pectin biosynthesis assay (MAPA)

V.E.1. Strategy. The basis of the MAPA is to rapidly determine if an expressed gene is a pectin biosynthetic GalAT, GlcAT, or KdoT. However, synthesis may require the presence of additional donor and acceptor substrates since pectin is a complex polymer and some of the selected genes may encode polymer backbone synthesizing enzymes. For example, a GalAT that is specific for xylogalacturonan (XGA) synthesis may require the presence of UDP-Xyl in the reaction while RG-I backbone synthesis will likely require both UDP-GalA and UDP-Rha. Thus, our strategy for the MAPA is to initially test each heterologously expressed gene in a minimum number of reactions that contain a mixture of substrates and acceptors so as to maximize the likelihood of recovering enzyme activity. Our initial strategy will be to test each expressed protein in four reactions that each contain a single radiolabeled nucleotide-sugar, a mixture of non-labeled nucleotide-sugars, and a mixture of acceptors (see Table II). However, we anticipate that as we identify more pectin biosynthetic enzymes and gain a better understanding of their requirements, the specifics of the number of reactions used in the initial screen, and the reaction components will change. The basic enzyme buffer, reaction components, and pH will be those developed for GalATs in general [7-10] and as adapted for *Arabidopsis GALAT1* (Sterling and Mohnen, unpublished). The conditions to be used are applicable to numerous plant wall glycosyltransferases[5]. The initial assay will be a precipitation assay developed by the PI [7] that should be broadly useful for pectin biosynthetic enyzmes. However, two other assays developed by the PI, a filter assay[87] and a non-radiolabeled PAGE assay[111] will be used if warrented. The PI and a Co-PI have extensive biochemical experience with multiple glycosyltransferases [7-10,31,112-114], pectin biosynthetic methyltransferases [16,17,115] and nucleotide-sugar biosynthetic enzymes [19,28,86,114,116].

If our predicted function of the selected putative pectin biosynthetic genes is correct, the selected genes require UDP-galacturonic (UDP-GalA), UDP-glucuronic acid (UDP-GlcA), CMP-Kdo, and UDP-rhamnose (UDP-Rha) as activated sugar substrates. Of these, only UDP-GlcA is commercially available. The remaining required sugar nucleotides will be synthesized by methods already established in our laboratories at the CCRC. The PI routinely synthesize UDP-[^{14}C]GalA[7,8,10,19] either by the 4-epimerization of UDP-[^{14}C]GlcA[19] or oxidation of UDP-[^{14}C]Gal[117]. A Co-PI (Bar-Peled) has developed methods to enzymatically synthesize UDP-Rha[113,114] and more recently identified, cloned, and expressed recombinant plant genes (Harper and Bar-Peled, unpublished) in *E. coli* that allow facile enzymatic generation of TDP-Rha and UDP-Rha (Watts and Bar-Peled, unpublished). We have produced CMP-Kdo by cloning and expressing two recombinant enzymes that convert CTP, PEP and Ara to CMP-Kdo (Hogans and Bar-Peled, unpublished). The PI is director of CarboSource Services, a part of the NSF-funded Plant Cell Wall Biosynthesis Research Network which produces and distributes UDP-Xyl (see section VII).

V.E.2. Multiple acceptor pectin biosynthesis assay (MAPA). Heterologously expressed recombinant proteins will be assayed for pectin biosynthetic activity by determining whether they catalyze the transfer of a radiolabeled sugar residue to any one of several pools of pectic acceptors (Table II). Acceptor preparations will be generated by treatment of homogalacturonan, RG-I and RG-II that we routinely isolate[14,15] with mild acid, hydrolytic enzymes, or combinations of these reagents (Table II). These treatments selectively cleave specific glycosidic bonds, thereby exposing potential acceptor sites to which glycosyl transferases might transfer a glycosyl residue. Monosaccharides and oligosaccharides (DP < 5) will be removed from the resulting mixtures by gel-permeation and/or ion-exchange chromatography. Some of the pectic polymers obtained will be further fragmented. For example, RG-I (Table II, prep 6) will be treated with arabinosidase/arabinanase. Ion-exchange chromatography will be

used to separate the partially degraded arabinan sidechains (prep 8) from the dearabinosylated RG-I (prep 7), which will be treated with RG-hydrolase to cleave the backbone, thereby generating dearabinosylated RG-I fragments (prep 12).

The structural features of each acceptor substrate preparation will be characterized by standard methods that have previously been developed and applied in our laboratory[14,15,20,118,119]. For example, the molecular weight distribution will be determined by SEC. Glycosyl residue compositions will be determined in triplicate by GC-MS analyses of alditol acetate (neutral residues) and TMS methyl glycoside (acidic + neutral residues) derivatives. Glycosyl linkage compositions will be performed by GC-MS analysis of partially methylated alditol acetate (PMAA) derivatives. The ratio of 2-linked to terminal GlcA and the ratio of 4-linked to terminal GalA will be determined by chemically reducing uronic acid residues to their 6,6-dideuterio hexosyl derivatives, which will be converted to PMAAs that are readily identified by GC-MS.

Table II. Acceptors to be used for the multiple acceptor pectin biosynthesis assay (MAPA). Boxes indicate the four acceptor pools that will be used in the initial screening assays. For example, Pool 1 will contain Preps 1-4.

Pool	Prep	Source	Treatment	Major Component Produced for use as an Acceptor Substrate
1	1	Commercial polygalacturonic acid	None	Polygalacturonic acid
	2	Commercial pectin	None	Methyl-esterified HG
	3	Commercial polygalacturonic acid	*Endo*-polygalacturonase (partial)	Oligogalacturonides (OGAs) (DP 3-23)
	4	Commercial pectin	*Endo*-polygalacturonase, pectin methylesterase, (partial)	Randomly methyl-esterified oligogalacturonides (DP 3-23)
2	5	*A. thaliana* seed mucilage	No treatment	RG-I backbone
	6	Cultured plant cells and tissues	*Endo*-polygalacturonase and SEC	Highly branched, native RG-I
	7	RG-I from Prep 6	Arabinanase (or mild acid) and ion-exchange	Dearabinosylated RG-I
	8	RG-I from Prep 6	Arabinanase (or mild acid) and ion-exchange	Partially degraded RG-I arabinan sidechains
	9	RG-I from Prep 6	*Endo*-galactanase and ion-exchange	Degalactosylated RG-I
	10	RG-I from Prep 6	*Endo*-galactanase and ion-exchange	Partially degraded RG-I galactan sidechains
3	11	RG-I backbone from Prep 5	RG Hydrolase +/- rhamnosidase[a]	RG-I backbone fragments
	12	Dearabinosylated RG-I from Prep 7	RG Hydrolase +/- rhamnosidase[a]	RG-I fragments[a] depleted in arabinan sidechains
	13	Degalactosylated RG-I from Prep 9	RG Hydrolase +/- rhamnosidase[a]	RG-I fragments[a] depleted in galactan sidechains
4	14	Cultured plant cells and tissues	Endo-polygalacturonase and SEC	RG-II
	15	RG-II from Prep 14	Mild acid	Partially fragmented RG-II

[a]RG-hydrolase cleaves the RG-I backbone, generating fragments with a GalA residue at the reducing terminus. In the absence of rhamnosidase, the non-reducing terminus of each fragment will be rhamnosyl, but in the presence of rhamnosidase, the non-reducing terminus will be GalA

Each acceptor substrate pool (Table II) will be mixed with a radiolabeled nucleotide sugar, a mixture of non-labeled nucleotide-sugars, and an aliquot of the recombinant gene product. Following incubation, acceptor and reaction products (including those that have been radiolabeled by the action of

an expressed glycosyltransferase) will be precipitated with methanol:chloroform[7,8] and non-incorporated nucleotide-sugar will be removed by repeated washing with aq 65% ethanol[7,8]. Glycosyltransferase activity will be detected by the presence of radioactivity in the pellet. The nature of the glycosyltransferase reaction catalyzed by a given enzyme will be deduced by examining increases in alcohol-insoluble radioactivity for different donor and acceptor combinations. An enzyme that transfers a glycosyl residue to a pectic polymer that has structural features common to more than one of the acceptor substrate pools will produce an increase in alcohol-insoluble radioactivity in more than one of the reaction mixtures. As the overall structural features of the components of each substrate pool will be known, this will provide information regarding the nature of glycosyltransferase activity, which can then be further characterized by performing glycosyl transferase assays using individual relevant acceptors/acceptor pools. Ultimately, highly purified acceptor substrates will be used and the radioactive products will be chemically characterized.

V.F. Characterization of putative pectin biosynthetic mutants

V.F.1. Strategy. Nineteen of our selected genes have one or more sequence-indexed T-DNA insertion mutants currently listed in the Salk Institute Genomic Analysis Laboratory (SIGnAL) *Arabidopsis* T-DNA mutant collection (http://signal.salk.edu/cgi-bin/tdnaexpress). We expect that additional T-DNA mutant lines will become available prior to, and during the time the research proposed is carried out, since the *Arabidopsis* sequence-indexed T-DNA insertion Project is only beginning its second year of funding. We will obtain these mutants as they are deposited. We have already received, grown, and selfed eight of the T-DNA mutants and the additional available mutants have been ordered from ABRC.

The putative pectin biosynthesis mutants will aid in the identification of gene function in two ways. The visible phenotypes of such mutants may provide information on the biological function of the gene (if there is no redundancy in gene function) by demonstrating when during growth and development the particular gene product is needed. Structural analysis of the pectin in the mutant walls may provide information about the specific enzyme activity of the gene in pectin synthesis. Furthermore, the mutants can also provide precursor polysaccharides that could be used as exogenous acceptors to identify enzyme activity. For these reasons, the available putative pectin biosynthetic mutants will be characterized for plant phenotype and wall structure and, when applicable, the pectic polysaccharides from such mutants will be isolated for use as exogenous acceptors. Finally, when applicable, particulate and soluble fractions from the mutants will be used for *in vitro* complementation enzyme assays to confirm the enzyme activity of the expressed gene.

V.F.2. Initial characterization of mutant phenotypes and bulking up of seed. A portion of the initial mutant seed obtained from ABRC (segregating T3 line, see http://signal.salk.edu/tdna_FAQs.html) will be grown and allowed to self-cross to increase the seed stock (T4). Multiple plants from T4 seed of each line will be grown and the presence of the T-DNA insert determined by PCR of plant genomic DNA using a T-DNA primer and a gene specific primer. The same DNA will be analyzed with gene specific primers that should span the T-DNA insertion site. These analyses should indicate whether the given plant contains a T-DNA insert and if so, whether it is homozygous or heterozygous for the mutation. If necessary, we will use Southern blotting and hybridization with the specific genes, as recommended by the SIGnAL web site, to determine if the gene contains the expected T-DNA insert.

Seed homozygous for the T-DNA insertion (when not lethal) or heterozygous (when no viable TDNA homozygous plants are obtained) will be selfed to amplify the seed and, for heterozygous plants, to test for segregation of any phenotype or T-DNA insert. If the parent plant was heterozygous for the T-DNA mutant, we would expect approximately a 3:1 mutant segregation in the next generation. If it was homozygous, we should see a 100% mutant phenotype. Plants will be scored as heterozygous or homozygous by PCR analysis of the T-DNA insert and by any visible phenotype. Homozygous or heterozygous plants will be used for growth phenotype and cell wall analysis. The seed will also be crossed with wild type Columbia and then selfed to eliminate the possibility that the lines contain an

unexpected mutation or additional T-DNA insert(s). All growth of mutant and wild type plants will be under controlled temperature, humidity and light conditions in growth chambers (Russell Malmberg, Plant Biology Dept., has agreed to advise us on the characterization of the mutants, see letter of suppport).

V.F.3. Growth Phenotype analysis. Several growth parameters of the mutant and wild type plants will be recorded to yield a general phenotypic characterization of the mutant plants. We will limit our initial phenotypic analyses to soil grown plants and use the growth stage designations of Boyes et al.[120]. Bi-weekly measurements will be taken of seed germination, cotyledon and hypocotyl emergence, numbe of rosette leaves >1 mm, plant height, inflorescence emergence, number of flowers, number of flower stalks, number of siliques, number and length of siliques at end of flowering stage, and number (mass) of seeds obtained.

V.F.4. Analysis of Cell Walls. Homozygous or heterozygous plants will also be grown and analyzed for wall composition and linkage. Cell walls will be prepared as alcohol insoluble residues. (AIRs) from WT and (homozygous) mutant *Arabidopsis* plant tissues[121]. AIRs will be prepared by homogenizing leaves and stems (from soil-grown plants) and roots (from liquid-cultured plants) in aq. 80% EtOH followed by washes with abs EtOH, chloroform-methanol, and acetone. Separate fractions containing RG-I, RG-II and oligogalacturonides will be obtained by size-exclusion chromatography (SEC) and ion exchange chromatography of the material solubilized from the cell walls by treatment with pectin methyl esterase (PME) and *endo*-polygalacturonase (EPG). The yields, glycosyl residue compositions, and glycosyl linkage compositions of each fraction will be determined[14].

V.G. Production and characterization of putative acceptors from mutant walls .

Cell walls and pectin subfractions containing RG-I, RG-II and homogalacturonans will be prepared as described above in section V.F.4. Specific pectic subfractions that will be used as acceptor substrates will be prepared and characterized as described in section V.E.2. (see Table IV).

V.H. Permeabilized membrane Assay

If no enzyme activity is discernible in the MAPA, membranes from wild type and mutant plants will be isolated, suspended in reaction buffer containing 0.1% Triton X-100 to permeabilize the membranes[10,30-32], and incubated in the presence and absence of both the heterologously expressed protein and the specific antibody against the protein. Inhibition of the transfer of a radiolabeled sugar (e.g. from UDP-[^{14}C]GalA) onto endogenous acceptor in membranes incubated with the antibody, compared to activity in the absence of the antibody, would indicate that the protein encoded a pectin biosynthetic enzyme specific for the endogenous acceptor. Recovery of activity in mutant permeabilized membranes to which the heterologously expressed wild-type protein was added (i.e. *in vitro* reconstitution) would confirm this activity. Pectic acceptor molecules would then be generated from the mutant plant cell walls and used to attempt to obtain activity in the MAPA.

V.I. Generation of antibodies

Polyclonal antibodies will be generated in two ways to enhance the likelihood of obtaining useful specific antibodies for each gene product. Our main strategy will be to use purified recombinantly expressed proteins to generate rabbit polyclonal antibodies at the University of Georgia Polyclonal Antibody Production Service. The recombinant proteins used for immunizations will lack any putative transmembrane domains and will be generated as described in section IV.B. Antisera that are obtained will be cleared of antibodies against the epitope tag sequence(s) incorporated into the heterologously expressed proteins by immunoabsorption against synthetic epitope tag peptides (to be synthesized at Molecular Genetics Instrumentation Facility, University of Georgia) coupled to a solid support.

The specificity of the antisera will be tested by immunoblotting against the full series of purified heterologously expressed proteins. In addition, the antibodies will be tested for their abilities to specifically bind to native proteins as assayed by immunoassays against extracts prepared from

Arabidopsis tissues. We will also test the antisera for their abilities to inhibit the activity of any enzymes for which *in vitro* activity has been demonstrated.

Given that the proteins within the different families being studied have domains showing strong sequence similarities/identities (e.g., the GalAT superfamily, see Figure 7), immunizing with the complete heterologously expressed proteins may not yield mono-specific sera. Thus, we will also immunize rabbits with an eight-branched, multiple antigenic protein (MAP)[122] synthesized based on peptide sequences derived those portions of the proteins that show little if any sequence conservation amongst the members of a given protein family (e.g, the *N*-terminal portion of the C-terminal domains of the GalAT superfamily). Colleagues at the CCRC have successfully used this approach to generate specific antisera against *Cochliobolus* endopolygalacturonase[123]. The MAPs will be synthesized at the Molecular Genetics Instrumentation Facility, University of Georgia, using the 8-branch Fmoc resin[123]. The strength of the MAP system is its ability to generate specific anti-peptide antibodies, but the ability of the MAP-generated antibodies to bind to intact proteins may be more limited. Thus, it may prove necessary to make multiple MAPs based on different non-conserved peptides and determine which, if any, of them recognize the complete proteins.

V.J. Immunocytochemistry: Subcellular localization of putative pectin biosynthetic enzymes

We propose to use immunocytochemistry to determine the subcellular localization of each of the pectin biosynthetic proteins within *Arabidopsis* cells from a tissue shown by RT-PCR to express the specific protein of interest. Such localizations will provide additional information with which to determine the role(s), if any, of the various polypeptides in pectin biosynthesis. For example, localization of a polypeptide to Golgi cisternae would be consistent with a function of that polypeptide in pectin biosynthesis synthesis. Localization of a polypeptide to another cellular compartment (e.g., the vacuole) would indicate that the polypeptide is unlikely to have a role in pectin biosynthesis. In the latter case, the polypeptide and corresponding gene will be given a lower priority for the further detailed studies that are outlined elsewhere in this application.

Both developing *Arabidopsis* seedlings and growing suspension cultures will be used, since both tissues contain actively dividing and elongating cells that are synthesizing walls. We have extensive experience with the preparation and use of these tissues for immunocytochemical studies both at the cellular and sub-cellular levels[39,40,124-127]. Tissue sections will be prepared from four day old *Arabidopsis* seedlings, or exponentially growing suspension cultured *Arabidopsis* cells, and immunolabeled as described[39]. Immunolabeled sections will be examined at 80kV with a Zeiss EM 902A Electron microscope available for our use in the Plant Biology Department at the University of Georgia.

VI. Broader impact of results including how project will promote teaching and training and participation of underrepresented groups

The PI and Co-PI's routinely advise multiple undergraduate students in undergraduate researcher and will continue to do so through the research described in this proposal . For example, the PI has mentored 30 undergraduate students over the last 12 years. Seven of these students were minority students. It is expected that each PI and Co-PI will mentor at least one undergraduate student in independent research on the proposed project for the duration of the proposed research. Since most biology-based majors at UGA require at least one semester of independent research, and since the Georgia Hope Scholarship program pays full tuition for any Georgian student who maintains a B average in college, there is a vast pool of bright and talented students who wish to carry out undergraduate research. Although we have only budgeted for 5 undergraduate students/yr, we anticipate training more students since many students do one or more semesters of undergraduate research for academic credit.

To increase the participation of historically underrepresented groups in the research pool, we have also arranged with Curtis D. Byrd (Director of minority Recruitment and Retention, University of Georgia) to recruit from the pool of outstanding undergraduates from historically underrepresented groups that come to UGA each summer through the Summer Undergraduate Research Program (SURP) (www.gradsch.uga.edu/rr/). The students will work with the PI, Co-PIs, graduate students and/or postdocs to carry out intensive summer research. The SURP students will also be expected to write a paper and present a poster describing their research project at the end of the research experience.

We have also arranged with Curtis D. Byrd to recruit one or more incoming minority graduate students into the project through the newly developed SURP-Bridge program. This program funds accepted minority graduate students the summer before their graduate work to carry out independent research at UGA and develop research skills so as to increase their likelihood of success in graduate school (see attached letter from Curtis D. Byrd).

VII. Results from Prior NSF Support

(Mohnen) NSF Resource Coordination Network (RCN) Grant No. 0090281; Ken Keegstra (PI), Debra Mohnen (Co-PI); 2/15/01-2/14/06; $559,867 total grant, $239,991; *Plant Cell Wall Biosynthesis Research Network*. One of the goals of the grant, that is carried out by the PI at the CRRC, is to coordinate and partially support the acquisition of substrates and acceptors for enzymes involved in plant cell wall polysaccharide biosynthesis. To do this, we established "CarboSource Services" (http://www.ccrc.uga.edu/web/services/carbosource/index.html) and a Resource Coordinator was hired who is responsible for producing rare substrates and making them available to the plant cell wall biosynthesis community. We have produced over 400 mg of UDP-xylose and distributed to 24 research groups throughout the world. We are also developing methods for the large scale production of UDP-GalA and the production of UDP-Ara for eventual distribution. Other accomplishments, including hosting the international Cell Wall Biosynthesis Meeting, May 12 - May15, 2002, at the UCLA Lake Arrowhead Conference Center are outlined on the Plant Cell Wall Biosynthesis Research Network are outlined at the website (http://xyloglucan.prl.msu.edu/).

(York) NSF Grant MCB-9974673; 08/01/99-07/31/03; $330,000; *Dynamic models of xyloglucan conformation in the primary cell walls of plants*. The major goal is to characterize the molecular and topological processes governing the interaction of xyloglucan with cellulose to form a load-bearing network in the primary cell wall. We performed experiments that lay the groundwork for this technically challenging, multidisciplinary project. We developed methods to prepare ^{13}C-enriched plant cell wall polysaccharides, in order to increase the sensitivity of the heteronuclear NMR experiments used for their conformational analysis. Onion plants were grown in the presence of ^{13}C-enriched CO_2 in a chamber specifically designed for this purpose. ^{13}C-Enriched xyloglucan and cello-oligomers were prepared from the cell walls of these plants. We established chemoenzymatic routes for the generation of several model oligoglycosides, including ^{13}C-enriched, β-methyl glycosides derived from cellulose and xyloglucan. These molecules are being used to assess the possibility of using "residual dipolar coupling" as a means of determining the average geometry and dynamic properties of these model compounds. We also extended computational methods to simulate the conformational dynamics of xyloglucan and performed molecular dynamics simulations for several of the model oligoglycosides. The results indicate that, in solution, xyloglucan and cellulose oligoglycosides have a net right-hand twist, in contrast to the flat conformation characteristic of cellulose and proposed for xyloglucan bound to its surface. Rigorous experimental evaluation of this result is being performed by NMR analysis of the model oligoglycosides and manuscripts describing these results are in preparation.

(Hahn) NSF Grant PGR-9772664; 12/01/98-11/30/02, *Medicago truncatula as the nodal species for comparative and functional legume genomics*. Total award: $3,416,562 (D. Cook, PI); Georgia sub-contract: $214,613. This was a multi-investigator, multi-institutional grant to establish *Medicago truncatula* as a model species for legume biology and combined structural and comparative genome analysis with functional genomics. Our laboratory participated in the functional genomics part of this research program, which was largely focused on the generation of a large and diverse expressed sequence tag (EST) collection from microbe-induced and developmentally-specific cDNA libraries. Specifically, we generated cDNA libraries from oligoglucoside elicitor- and oligogalacturonide elicitor-treated root tissues of *M. truncatula*. A total of 2596 5'-ESTs from the oligoglucoside elicitor library and 2861 5'-ESTs from the oligogalacturonide library were sequenced and deposited in GenBank. In addition, we generated a subtraction cDNA library using *M. truncatula* roots infected with different strains of *Sinorhizobium meliloti*. The two *S. meliloti* strains chosen vary in their ability to nodulate *M. truncatula* ecotypes: *S. meliloti* NR 185 induces normal elongated pink nodules on *M. truncatula* ecotype A17, while *S. meliloti* Rm41 produces only white round nodules on this ecotype. The phenotypes of these two *S. meliloti* strains are reversed on *M. truncatula* ecotype A20. About 1000 clones from the subtraction library were sequenced from both ends and are being prepared for submission to GenBank. These sequences are being analyzed for unique sequences not otherwise represented in the *M. truncatula* EST collection and for those that are differentially regulated in the *S. meliloti-M. truncatula* interaction. Publications are being prepared summarizing the results of the EST sequencing effort.

References

1. H. Iwai, N. Masaoka, T. Ishii, and S. Satoh, A pectin glucuronyltransferase gene is essential for intercellular attachment in the plant meristem, Proc. Natl. Acad. Sci. USA 99:16319-16324 (2002).
2. K.A. White, I.A. Kaltashov, R.J. Cotter, and C.R.H. Raetz, A Mono-functional 3-deoxy-D-manno-octulosonic acid (Kdo) transferase and a Kdo kinase in extracts of *Haemophilus influenzae*, The J. Biol. Chem. 272:16555-16563 (1997).
3. Z.-Y. Wang, F.-Q. Zheng, G.-Z. Shenm, J.-P. Gao, D.P. Snustad, M.-G. Li, J.-L. Zhang, and M.-H. Hong, The amylose content in rice endosperm is related to the post-transcriptional regulation of the *waxy* gene, Plant J. 7:613-622 (1995).
4. T. Clementz and C.R.H. Raetz, A gene coding for 3-deoxy-D-*manno*-octulosonic-acid transferase in *Eschericial coli.* Identification, mapping, cloning, and sequencing, J. Biol. Chem. 266:9687-9696 (1991).
5. D. Mohnen, Biosynthesis of pectins, in: "Pectins and their Manipulation", G.B. Seymour et al., Blackwell Publishing and CRC Press, Oxford, pp. 52-98 (2002).
6. V. Walbot, Arabidopsis thaliana genome. A green chapter in the book of life, Nature 408:794-795 (2000).
7. R.L. Doong, K. Liljebjelke, G. Fralish, A. Kumar, and D. Mohnen, Cell free synthesis of pectin: identification and partial characterization of polygalacturonate 4-α-galacturonosyltransferase and its products from membrane preparations of tobacco (*Nicotiana tabacum* L. cv samsun) cell suspension cultures, Plant Physiol. 109:141-152 (1995).
8. R.L. Doong and D. Mohnen, Solubilization and characterization of a galacturonosyltransferase that synthesizes the pectic polysaccharide homogalacturonan, Plant J. 13:363-374 (1998).
9. H.V. Scheller, R.L. Doong, B.L. Ridley, and D. Mohnen, Pectin biosynthesis: a solubilized galacturonosyltransferase from tobacco catalyzes the transfer of galacturonic acid from UDP-galacturonic acid onto the non-reducing end of homogalacturonan, Planta 207:512-517 (1999).
10. J. Sterling, H.F. Quigley, A. Orellana, and D. Mohnen, The catalytic site of the pectin biosynthetic enzyme α-1,4-galacturonosyltransferase (GalAT) is located in the lumen of the Golgi, Plant Physiol. 127:360-371 (2001).
11. B.L. Ridley, M.A. O'Neill, and D. Mohnen, Pectins: structure, biosynthesis, and oligogalacturonide-related signaling, Phytochemistry 57:929-967 (2001).
12. M.A. O'Neill, D. Warrenfeltz, K. Kates, P. Pellerin, T. Doco, A.G. Darvill, and P. Albersheim, Rhamnogalacturonan-II, a pectic polysaccharide in the walls of growing plant cell, forms a dimer that is covalently cross-linked by a borate ester - in vitro conditions for the formation and hydrolysis of the dimer, J. Biol. Chem. 271:22923-22930 (1996).
13. E. Zablackis, W.S. York, M. Pauly, S. Hantus, W.-D. Reiter, C.C.S. Chapple, P. Albersheim, and A. Darvill, Substitution of L-fucose by L-galactose in cell walls of arabidopsis *mur*1, Science 272:1808-1810 (1996).
14. M. O'Neill, P. Albersheim, and A. Darvill, The pectic polysaccharides of primary cell walls, in: "Methods in Plant Biochemistry, Volume 2", P.M. Dey., ed., Academic Press, London, pp. 415-441 (1990).

15. W.S. York, A.G. Darvill, M. McNeil, T.T. Stevenson, and P. Albersheim, Isolation and characterization of plant cell walls and cell wall components. Methods Enzymol. 118:3-40 (1985).
16. F. Goubet and D. Mohnen, Subcellular localization and topology of homogalacturonan methyltransferase in suspension-cultured *Nicotiano tabacum* cells, Planta 209:112-117 (1999).
17. F. Goubet and D. Mohnen, Solubilization and partial characterization of homogalacturonan-methyltransferase from microsomal membranes of suspension-cultured tobacco cells, Plant Physiol. 121:281-290 (1999).
18. D. Mohnen, Biosynthesis of pectins and galactomannans, in: "Comprehensive Natural Products Chemistry, Vol. 3, Carbohydrates and Their Derivatives including Tannins, Cellulose, and Related Lignins", B.M. Pinto., ed., Elsevier, Oxford, pp. 497-527 (1999).
19. K. Liljebjelke, R. Adolphson, K. Baker, R.L. Doong, and D. Mohnen, Enzymatic synthesis and purification of uridine diphosphate [^{14}C]galacturonic acid: a substrate for pectin biosynthesis, Anal. Biochem. 225:296-304 (1995).
20. M.A. O'Neill, S. Eberhard, P. Albersheim, and A.G. Darvill, Requirement of borate cross-linking of cell wall rhamnogalacturonan II for *Arabidopsis* growth, Science 846-849 (2001).
21. M. Pauly, Q. Qin, H. Greene, P. Albersheim, A. Darvill, and W.S. York, Changes in the structure of xyloglucan during cell elongation, Planta 212:842-850 (2001).
22. D. Mohnen and M.G. Hahn, Cell wall carbohydrates as signals in plants, Sem. Cell Biol. 4:93-102 (1993).
23. S. Eberhard, N. Doubrava, V. Marfà, D. Mohnen, A. Southwick, A. Darvill, and P. Albersheim, Pectic cell wall fragments regulate tobacco thin-cell-layer explant morphogenesis, Plant Cell 1:747-755 (1989).
24. V. Marfà, D.J. Gollin, S. Eberhard, D. Mohnen, A. Darvill, and P. Albersheim, Oligogalacturonides are able to induce flowers to form on tobacco explants. Plant J. 1:217-225 (1991).
25. M.G. Hahn, A.G. Darvill, and P. Albersheim, Host-pathogen interactions XIX. The endogenous elicitor, a fragment of a plant cell wall polysaccharide that elicits phytoalexin accumulation in soybeans, Plant Physiol. 68:1161-1169 (1981).
26. F. Cervone, M.G. Hahn, G. De Lorenzo, A. Darvill, and P. Albersheim, Host-pathogen interactions XXXIII. A plant protein converts a fungal pathogenesis factor into an elicitor of plant defense responses, Plant Physiol. 90:542-548 (1989).
27. G. De Lorenzo, F. Cervone, M.G. Hahn, A. Darvill, and P. Albersheim, Bacterial endopectate lyase: evidence that plant cell wall pH prevents tissue maceration and increases the half-life of elicitor-active oligogalacturonides, Physiol. Mol. Plant Pathol. 39:335-344 (1991).
28. M. Bar-Peled, C.L. Griffith, and T.L. Doering, Functional cloning and characterization of a UDP-glucuronic acid decarboxylase: the pathogenic fungus *Cryptococcus neoformans* elucidates UDP-xylose synthesis, Proc. Natl. Acad. Sci. USA 98:12003-12008 (2001).
29. S.U. Ahmed, M. Bar-Peled, and N.V. Raikhel, Clonging and subcellular location of an arabidopsis receptor-like protein that share common features with proteins- sorting receptors of eukaryotic cells, Plant Physiol. 114:325-336 (2000).
30. M. Bar-Peled and N.V. Raikhel, Characterization of AtSEC12 and AtSAR1 - proteins likely involved in endoplasmic reticulum and Golgi transport, Plant Physiol. 114:315-324 (2000).

31. R.M. Perrin, A.E. DeRocher, M. Bar-Peled, W. Zeng, L. Norambuena, A. Orellana, N.V. Raikhel, and K. Keegstra, Xyloglucan fucosyltransferase, an enzyme involved in plant cell wall biosynthesis, Science 284:1976-1979 (1999).
32. S.U. Ahmed, M. Bar-Peled, and N.V. Raikhel, Cloning and subcellular location of an arabidopsis receptor-like protein that shares common features with protein-sorting receptors of eukaryotic cells. Plant Physiol. 115:311-312 (1997).
33. M. Bar-Peled and N.V. Raikhel, Characterization of ATSEC12 and ATSAR1-proteins likely involved in the endoplasmic-reticulum and Golgi transport, Plant Physiol. 114:315-324 (1997).
34. M. Bar-Peled and N.V. Raikhel, An efficient method for cloning in-frame fusion protein genes, Anal. Biochem. 250:262-264 (1997).
35. M. Bar-Peled, D.C. Bassham, and N.V. Raikhel, Transport of proteins in eukaryotic cell-more questions ahead, Plant Mol. Biol. 32:223-249 (1996).
36. M. Bar-Peled, A. da Silva Conceicao, L. Frigerio, and N.V. Raikhel, Expression and regulation of aERD2, a gene encoding the KDEL receptor homolog in plants, and other genes encoding proteins involved in ER-Golgi vesicular trafficking, Plant Cell 7:667-676 (1995).
37. M. Bar-Peled and N.V. Raikhel, A method for isolation and purification of specific antibodies to a protein fused in the GST. Anal. Biochem. 241:140-142 (1996).
38. R. Balestrini, M.G. Hahn, A. Faccio, K. Mendgen, and P. Bonfante, Differential localization of carbohydrate epitopes in plant cell walls in the presence and absence of arbuscular mycorrhizal fungi, Plant Physiol. 111:203-213 (1996).
39. G. Freshour, R.P. Clay, M.S. Fuller, P. Albersheim, A.G. Darvill, and M.G. Hahn, Developmental and tissue-specific structural alterations of the cell-wall polysaccharides of *Arabidopsis thaliana* Roots, Plant Physiol. 110:1413-1429 (1996).
40. M.N.V. Williams, G. Freshour, A.G. Darvill, P. Albersheim, and M.G. Hahn, An antibody Fab selected from a recombinant phage display library detects deesterified pectic polysaccharide rhamnogalacturonan II in plant cells, Plant Cell 8:673-685 (1996).
41. G. Kossoy, H. Ben-Hur, I. Zusman, and Z. Madar, Effects of a 15% orange-pulp diet on tumorigenesis and immune response in rats with colon tumors, Oncol Rep 8:1387-1391 (2001).
42. I. Eliaz, The potential role of modified citrus pectin in the prevention of cancer metastasis, Clinical Practice of Alternative Medicine 2:1-7 (2001).
43. D. Platt and A. Raz, Modulation of the lung colonization of B16-F1 melanoma cells by citrus pectin, J. Natl. Cancer Inst. 84:438-442 (1992).
44. K.J. Pienta, H. Naik, A. Akhtar, K. Yamazaki, T.S. Replogle, J. Lehr, T.L. Donat, L. Tait, V. Hogan, and A. Raz, Inhibition of spontaneous metastasis in a rat prostate cancer model by oral administration of modified citrus pectin, J. Natl. Cancer Inst. 87:348-353 (1995).
45. J.F. Pirani, The effects of phytotherapeutic agents on prostate cancer: an overview of recent clinical trials of PC SPES, Urology 58:36-38 (2001).
46. H.A. Schols, M.A. Posthumus, and A.G.J. Voragen, Structural features of hairy regions of pectins isolated from apple juice produced by the liquefaction process, Carbohydr. Res. 206:117-129 (1990).
47. H.A. Schols, E.J. Bakx, D. Schipper, and A.G.J. Voragen, A xylogalacturonan subunit present in the modified hairy regions of apple pectin, Carbohydr. Res. 279:265-279 (1995).

48. L. Yu and A.J. Mort, Partial characterization of xylogalacturonans from cell walls of ripe watermelon fruit: Inhibition of endopolygalacturonase activity by xylosylation, in: "Pectins and Pectinases", J. Visser et al., Elsevier, Amsterdam, pp. 79-88 (1996).
49. D.A. Hart and P.K. Kindel, Isolation and partial characterization of apiogalacturonans from the cell wall of *Lemna minor*, Biochem. J. 116:569-579 (1970).
50. R.R. Watson and N.S. Orenstein, Chemistry and biochemistry of apiose, Adv. Carbohydr. Chem. Biochem. 31:135-184 (1975).
51. A. Fleischer, M.A. O'Neil, and R. Ehwald, The pore size of non-graminaceous plant cell walls is rapidly decreased by borate ester cross-linking of the pectic polysaccharide rhamnogalacturonan II, Plant Physiol. 121:829-838 (1999).
52. J.-C. Mollet, S.-Y. Park, E.A. Nothnagel, and E.M. Lord, A lily stylar pectin is necessary for pollen tube adhesion to an in vitro stylar matrix, Plant Cell 12:1737-1749 (2000).
53. T.L. Western, J. Burn, W.L. Tan, D.J. Skinner, L. Martin-McCaffrey, B.A. Moffatt, and G.W. Haughn, Isolation and characterization of mutants defective in seed coat mucilage secretory cell development in Arabidopsis, Plant Physiol. 127:998-1011 (2001).
54. W.G.T. Willats, L. McCartney, and J.P. Knox, In-situ analysis of pectic polysaccharides in seed mucilage and at the root surface of *Arabidopsis thaliana*, Planta 213:37-44 (2001).
55. Z.H. González-Carranza, C.A. Whitelaw, R. Swarup, and J.A. Roberts, Temporal and spatial expression of a polygalacturonase during leaf and flower abscission in oilseed rape and Arabidopsis, Plant Physiol. 128:534-543 (2002).
56. M.A. Zwieniecki, P.J. Melcher, and N.M. Holbrook, Hydrogel control of xylem hydraulic resistance in plants, Science 291:1059-1062 (2001).
57. F. Côté, K.-S. Ham, M.G. Hahn, and C.W. Bergmann, Oligosaccharide elicitors in host-pathogen interactions generation, perception, and signal transduction, in: "Plant-Microbe Interactions", B.B. Biswas et al., Plenum Press, New York, pp. 385-432 (1998).
58. S. Fry, S. Aldington, P. Hetherington, and J. Aitken, Oligosaccharides as Signals and Substrates in the Plant Cell Wall, Plant Physiol. 103:1-5 (1993).
59. N. Shibuya and E. Minami, Oligosaccharide signalling for defence responses in plant, Physiol. Mol. Plant Pathol. 59:223-233 (2001).
60. R.G. Atkinson, R. Schröder, I.C. Hallett, D. Cohen, and E.A. MacRae, Overexpression of polygalacturonase in transgenic apple trees leads to a range of novel phenotypes involving changes in cell adhesion, Plant Physiol. 129:122-133 (2002).
61. M. Skjot, M. Pauly, M.S. Bush, B. Borkhardt, M.C. McCann, and P. Ulvskov, Direct interference with rhamnogalacturonan I biosynthesis in Golgi vesicles, Plant Physiol. 129:95-102 (2002).
62. S. Bouton, E. Leboeuf, and G. Mouille, QUASIMODO1 encodes a putative membrane-bound glycosyltransferase required for normal pectin synthesis and cell adhesion in Arabidopsis, Plant Cell 14:2577-2590 (2002).
63. D. Mohnen, R.L. Doong, K. Liljebjelke, G. Fralish, and J. Chan, Cell free synthesis of the pectic polysaccharide homogalacturonan, in: "Pectins and Pectinases", J. Visser et al., Elsevier Science B.V. Amsterdam, pp. 109-126 (1996).
64. N.C. Carpita and D.M. Gibeaut, Structural models of primary cell walls in flowering plants: consistency of molecular structure with the physical properties of the walls during growth, Plant J. 3:1-30 (1993).

65. J. An, M.A. O'Neill, P. Albersheim, and A.G. Darvill, Isolation and structural characterization of β-D-glucosyluronic acid and 4-*O*-methyl β-D-glucosyluronic acid-containing oligosaccharides from the cell-wall pectic polysaccharide, rhamnogalacturonan I, Carbohydr. Res. 252:235-243 (1994).
66. J. An, L. Zhang, M.A. O'Neill, P. Albersheim, and A.G. Darvill, Isolaton and structural characterization of endo-rhamnogalacturonase-generated fragments of the backbone of rhamnogalacturonan I, Carbohydr. Res. 264:83-96 (1994).
67. W.G.T. Willats, L. McCartney, W. Mackie, and J.P. Knox, Pectin: cell biology and prospects for functional analysis, Plant Mol. Biol. 47:9-27 (2001).
68. A. Kikuchi, Y. Edashige, T. Ishii, and S. Satoh, A xylogalacturonan whose level is dependent on the size of cell clusters is present in the pectin from cultured carrot cells, Planta 200:369-372 (1996).
69. G.O. Aspinall, Chemistry of cell wall polysaccharides. in: "The Biochemistry of Plants, Vol.3.", J. Preiss., ed., Academic Press, New York, pp. 473-500 (1980).
70. E.M. Lord and J.-C. Mollet, Plant cell adhesion: a bioassay facilitates discovery of the first pectin biosynthetic gene, Proc. Natl. Acad. Sci. USA 99:15843-15845 (2002).
71. D.H. Northcote, The Golgi apparatus. Endeavor 30:26-33 (1971).
72. D.H. Northcote and J.D. Pickett-Heaps, A function of the Golgi Apparatus in polysaccharide synthesis and transport in the root-cap cells of wheat, Biochem. J. 98:159-167 (1966).
73. P.J. Harris and D.H. Northcote, Polysaccharide formation in plant golgi bodies. Biochim. Biophys. Acta 237:56-64 (1971).
74. R.W. Stoddart and D.H. Northcote, Metabolic relationships of the isolated fractions of the pectic substances of actively growing sycamore cells. Biochem. J. 105:45-59 (1967).
75. G.F. Zhang and L.A. Staehelin, Functional compartmentation of the golgi apparatus of plant cells; Immunocytochemical analysis of high-pressure frozen- and freeze-substituted sycamore maple suspension culture cells, Plant Physiol. 99:1070-1083 (1992).
76. L.A. Staehelin and I. Moore, The plant Golgi apparatus: structure, functional organization and trafficking mechanisms, Annu. Rev. Plant Physiol. Plant Mol. Biol. 46:261-288 (1995).
77. M.E. Edwards, C.A. Dickson, S. Chengappa, C. Sidebottom, M.J. Gidley, and J.S.G. Reid, Molecular characterisation of a membrane-bound galactosyltransferase of plant cell wall matrix polysaccharide biosynthesis, Plant J. 19:691-697 (1999).
78. A. Faik, N.J. Price, N.V. Raikhel, and K. Keegstra, An *Arabidopsis* gene encoding an α-xylosyltransferase involved in xyloglucan biosynthesis, Proc. Natl. Acad. Sci. USA 99:7797-7802 (2002).
79. D.P. Delmer, A hot mutant for cellulose synthesis, Trends in Plant Science 3:164(1998).
80. C.L. Villemez, A.L. Swanson, and W.Z. Hassid, Properties of a polygalacturonic acid-synthesizing enzyme system from *Phaseolus aureus* seedlings. Arch. Biochem. Biophys. 116:446-452 (1966).
81. H. Kauss and A.L. Swanson, Cooperation of enzymes responsible for polymerization and methylation in pectin biosynthesis. Z. Naturforsch. 24:28-33 (1969).
82. T.-Y. Lin, A.D. Elbein, and J.C. Su, Substrate specificity in pectin synthesis. Biochem. Biophys. Res. Commun. 22:650-657 (1966).
83. G.P. Bolwell, G. Dalessandro, and D.H. Northcote, Decrease of polygalacturonic acid synthase during xylem differentiation in sycamore. Phytochemistry 24:699-702 (1985).

84. Y. Takeuchi and Y. Tsumuraya, In vitro biosynthesis of homogalacturonan by a membrane-bound galacturonosyltransferase from epicotyls of azuki bean, Biosci. Biotech. Biochem. 65:1519-1527 (2001).
85. K. Akita, T. Ishimizu, T. Tsukamoto, T. Ando, and S. Hase, Successive glycosyltransfer activity and enzymatic characterization of pectic polygalacturonate 4-α-galacturonosyltransferase solubilized from pollen tubes of *Petunia axillaris* using pyridylaminated oligogalacturonates as substrates, Plant Physiol. 130:374-379 (2002).
86. A. Orellana and D. Mohnen, Enzymatic synthesis and purification of [^{3}H] uridine diphosphate galacturonic acid for use in studying Golgi-localized transporters, Analytical Biochemistry 272:224-231 (1999).
87. J.A. Lemons, I.F. Forkner, J. Sterling, and D. Mohnen, Development of a filter assay for measuring homogalacturonan: α1,4-galacturonosyltransferase activity, in preparation.
88. S. Pagès, H.C.M. Kester, J. Visser, and J.A.E. Benen, Changing a single amino acid residue switches processive and non-processive behavior of *Aspergillus niger* endopolygalacturonase I and II, J. Biol. Chem. 276:33652-33656 (2001).
89. R.A.F. Reithmeier and C.M. Deber, Intrinsic membrane protein structure: principles and prediction, in: "The Structure of Biological Membranes", P. Yeagle., ed., CRC Press, Boca Raton, pp. 337-393 (1992).
90. J.A. Campbell, G.J. Davies, V. Bulone, and B. Henrissat, A classification of nucleotide-diphospho-sugar glycosyltransferases based on amino acid sequence similarities, Biochem. J. 326:929-942 (1997).
91. B. Henrissat and G.J. Davies, Glycoside hydrolases and glycosyltransferases. Families, modules, and implications for genomics. Plant Physiol 124:1515-1519 (2000).
92. R. Sarria, T.A. Wagner, M.A. O'Neill, A. Faik, C.G. Wilkerson, K. Keegstra, and N.V. Raikhel, Characterization of a family of Arabidopsis genes related to xyloglucan fucosyltransferase1, Plant Physiol. 127:1595-1606 (2001).
93. S.J. Charnock, B. Henrissat, and G.J. Davies, Three-dimensional structures of UDP-sugar glycosyltransferases illuminate the biosynthesis of plant polysaccharides, Plant Physiol. 125:527-531 (2001).
94. C. Whitfield and M.A. Valvano, Biosynthesis and expression of cell-surface polysaccharides in gram-negative bacteria, Adv. Microb. Physiol. 35:135-246 (1993).
95. M. Regué, N. Climent, N. Abitiu, N. Coderch, S. Merino, L. Izquierdo, M. Altarriba, and J.M. Tomás, Genetic characterization of the *Klebsiella pneumoniae waa* gene cluster, involved in core lipopolysaccharide biosynthesis, J. Bacteriol. 183:3564-3573 (2001).
96. J. Shang, R. Qiu, J. Wang, J. Liu, R. Zhou, H. Ding, S. Yang, S. Zhang, and C. Jin, Molecular cloning and expresssion of Gal1,3GalNAc 2,3-sialyltransferase from human fetal liver, Eur. J. Biochem. 265:580-588 (1999).
97. W. Gillespie, S. Kelm, and J.C. Paulson, Cloning and expression of the Gal beta 1,3GalNAc alpha 2,3-sialyltransferase, J. Biol. Chem. 267:21004-21010 (1992).
98. B.D. Livingstom and J.C. Paulson, Polymerase chain reaction cloning of a developmentally regulated member of the sialyltransferase gene family. J. Biol. Chem. 265:580-588 (1993).
99. D. Mohnen, H. Shinshi, G. Felix, and F.J. Meins, Hormonal regulation of β1,3-glucanase messenger RNA levels in cultured tobacco tissues. EMBO J. 4:1631-1635 (1985).

100. H. Shinshi, D. Mohnen, and F. Meins, Jr. Regulation of a plant pathogenesis-related enzyme: Inhibition of chitinase and chitinase mRNA accumulation in cultured tobacco tissues by auxin and cytokinin, Proc. Natl. Acad. Sci. USA 84:89-93 (1987).
101. C.M. Cabral, Y. Liu, K.W. Moremen, and R.N. Sifers, Organizational diversity among distinct glycoprotein endoplasmic reticulum-associated degradation programs, Mol. Biol. Cell 13:2639-2650 (2002).
102. A.C. Porchia and H.V. Scheller, Arabinoxylan biosynthesis: identification and partial characterization of beta-1,4-xylosytransferase from wheat, physiological plantarum 110:350-356 (2000).(Abstract)
103. J.F. Kukowska-Latallo, R.D. Larsen, R.P. Nair, and J.B. Lowe, A cloned human cDNA determines expression of a mouse stage-specific embryonic antigen and the Lewis blood group alpha(1,3/1,4)fucosyltransferase, Genes Dev. 4:1288-1303 (1990).
104. M. Mallisard, S. Zeng, and E.G. Berger, Expression of functional soluble forms of human • -1,4-galactosyltransferase I, • -2,6-sialyltransferase, and • -2,6-sialyltransferase, and • -1,3-fucosyltransferase VI in the methylotrophic yeast *Pichia pastoris* , Biochem. Biophys. Res. Commun. 267:169-173 (2000).
105. P.A. Romero, M. Lussier, A.M. Sdicu, H. Bussey, and A. Herscovics, KTR1P is an alpha-1,2-mannosyltransferase of *Saccharomyces cerevisiae* comparison of the enzymatic properties of soluble recombinant KTR1P and KRE2P/MNT1P produced in *Pichia pastoris*, Biochem. J. 321:289-295 (1997).
106. U. Hochstrasser, M. Luscher, C. DeVirgilio, T. Boller, and A. Wiemken, Expression of a functional barley sucrose-fructan 6-fructosyltransferase in the methylotrophic yeast *Pichia pastoris*, Febs Letters 440:356-360 (1998).
107. P.F. Gallet, H. Vaujour, J.-M. Petit, A. Maftah, A. Oulmouden, R. Oriol, C. Le Narvor, M. Guilloton, and R. Julien, Heterologous expression of an engineered truncated form of human Lewis fucosyltransferase (Fuc-TIII) by the methylotrophic yeast *Pichia pastoris*, Glycobiology 8:919-925 (1998).
108. K.W. Moreman and O. Touster, Topology of mannosidasw II in rat liver membrane and release of the caalytic domain by selective proteolysis, Jr Biol. Chem 23:10945-10951 (1986).
109. R.K. Merkle, Y. Zhang, P.J. Ruest, Y.F. Liao, and K.W. Moremen, Cloning, expression, purification, and charactarization of the murine lysosomal acid alpha-mannosidase, Biochim. Biophys. Acta 1336:132-146 (1997).
110. Y.F. Liao, A. Lal, and K.W. Moremen, Cloning, expression, purification, and characterization of the human broad-specificity lysosomal acid alpha-mannosidase, J. Biol. Chem. 271:28348-28358 (1996).
111. H.F. Quigley, D.L. Howard, J. Sterling, and D. Mohnen, A non-radioactive gel electrophoresis assay for homogalacturonan • -1,4-galacturonosyltransferase, Anal. Biochem. (2002).
112. A. Faik, M. Bar-Peled, A. DeRocher, W. Zeng, R.M. Perrin, C. Wilkinson, N. Raikhel, and K. Keegstra, Biochemical characterization and molecular cloning of an • -1,2-fucosyltransferase that catalyzes the last step of cell wall xyloglucan biosynthesis in pea, The Journal of Biological Chemistry 275:15082-15089 (2000).

113. M. Bar-Peled, R. Fluhr, and J. Gressel, Juvenile-specific localization and accumulation of a rhamnosyltransferase and its bitter flavonoid in foliage, flowers, and young citrus fruits, Plant Physiol. 103:1377-1384 (1993).
114. M. Bar-Peled, E. Lewinsohn, R. Fluhr, and J. Gressel, UDP-rhamnose:flavanone-7-*O*-glucoside-2"-*O*-rhamnosyltransferase. Purification and characterization of an enzyme catalyzing the production of bitter compounds in citrus, J. Biol. Chem. 266:20953-20959 (1991).
115. F. Goubet, L.N. Council, and D. Mohnen, Identification and partial characterization of the pectin methyltransferase "homogalacturonan-methyltransferase" from membranes of tobacco cell suspensions, Plant Physiol 116:337-347 (1998).
116. A.D. Harper and M. Bar-Peled, Biosynthesis of UDP-Xylose. Cloning and characterization of a novel *Arabidopsis* gene family, *UXS*, encoding soluble and putative membrane-bound UDP-Glucuronic acid decarboxylase isoforms, Plant Physiol. 130:2188-2198 (2002).
117. S.S. Basu, G.D. Dotson, and C.R.H. Raetz, A facile enzymatic synthesis of uridine diphospho-[14C]galacturonic acid, Anal. Biochem. 280:173-177 (2000).
118. S. Vidal, T. Doco, P. Williams, P. Pellerin, W.S. York, M.A. O'Neill, J. Glushka, A.G. Darvill, and P. Albersheim, Structural characterization of the pectic polysaccharide rhamnogalacturonan II: evidence for the backbone location of the aceric acid-containing oligoglycosyl side chain, Carbohydr. Res. 326:277-294 (2000).
119. A.J. Whitcombe, M.A. O'Neill, W. Steffan, P. Albersheim, and A.G. Darvill, Structural characterization of the pectic polysacchride, Rhamnogalacturonan-II, Carbohydr. Res. 271:15-29 (1995).
120. D.C. Boyes, A.M. Zayed, R. Ascenzi, A.J. McCaskill, N.E. Hoffman, K.R. Davis, and J. Görlach, Growth stage-based phenotypic analysis of arabidopsis: a model for high throughput functional genomics in plants, Plant Cell 13:1499-1510 (2001).
121. E. Zablackis, J. Huang, B. Müller, A.G. Darvill, and P. Albersheim, Characterization of the cell-wall polysaccharides of *Arabidopsis thaliana* leaves, Plant Physiol. 107:1129-1138 (1995).
122. J.P. Briand, C. Barin, M.H. van Regenmortel, and S. Muller, Application and limitations of the multiple antigen peptide (MAP) system in the production and evaluation of anti-peptide and anti-protein antibodies, J. Immunol. Methods 156:255-265 (1992).
123. V. Ziehl, R.P. Clay, C. Bergmann, A. Darvill, and P. Albersheim, Polyclonal antibodies directed against synthetic N-terminus of fungal endopolygalacturonase recognizes the native protein and cross-reacts with endopolygalacturonase of a different funal species, (2002). (UnPub)
124. G. DiLaurenzio, J. Wysocka-Diller, J.E. Malamy, L. Pysh, Y. Helariutta, G. Freshour, M.G. Hahn, K.A. Feldmann, and P.N. Benfey, The SCARECROW gene regulates an asymmetric cell division that is essential for generating the radial organization of the Arabidopsis root. Cell 86:423-433 (1996).
125. M.G. Hahn, G. Freshour, W.-D. Reiter, P. Albersheim, and A.G. Darvill, Immunolocalization of fucose-containing polysaccharides in cell walls of the mur1 mutant of *Arabidopsis thaliana*, Plant Physiol (2002) (in press).
126. K. Enkerli, M.G. Hahn, and C.W. Mims, Ultrastructure of compatible and incompatible interactions of soybean roots infected with the plant pathogenic comycete Phytophthora sojae, Can. J. Bot. 75:1493-1508 (2002).

127. K. Enkerli, M.G. Hahn, and C.W. Mims, Immunogold localization of callose and other plant cell wall components in soybean roots infected with the plant pathogenic oomycete Phytophthora sojae, Can. J. Bot. 75:1509-1517 (1997).
128. J.D. Thompson, T.J. Gibson, F. Plewniak, F. Jeanmougin, and D.G. Higgins, The ClustalX windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Res. 24:4876-4882 (1997)

The amino acids which occur in the various amino acid sequences referred to in the specification have their usual three- and one-letter abbreviations routinely used in the art: A, Ala, Alanine; C, Cys, Cysteine; D, Asp, Aspartic Acid; E, Glu, Glutamic Acid; F, Phe, Phenylalanine; G, Gly, Glycine; H, His, Histidine; I, Ile, Isoleucine; K, Lys, Lysine; L, Leu, Leucine; M, Met, Methionine; N, Asn, Asparagine; P, Pro, Proline; Q, Gln, Glutamine; R, Arg, Arginine; S, Ser, Serine; T, Thr, Threonine; V, Val, Valine; W, Try, Tryptophan; Y, Tyr, Tyrosine.

A protein is considered an isolated protein if it is a protein isolated from a host cell in which it is recombinantly produced. It can be purified or it can simply be free of other proteins and biological materials with which it is associated in nature.

An isolated nucleic acid is a nucleic acid the structure of which is not identical to that of any naturally occurring nucleic acid or to that of any fragment of a naturally occurring genomic nucleic acid spanning more than three separate genes. The term therefore covers, for example, (a) a DNA which has the sequence of part of a naturally occurring genomic DNA molecule but is not flanked by both of the coding or noncoding sequences that flank that part of the molecule in the genome of the organism in which it naturally occurs; (b) a nucleic acid incorporated into a vector or into the genomic DNA of a prokaryote or eukaryote in a manner such that the resulting molecule is not identical to any naturally occurring vector or genomic DNA; (c) a separate molecule such as a cDNA, a genomic fragment, a fragment produced by polymerase chain reaction (PCR), or a restriction fragment; and (d) a recombinant nucleotide sequence that is part of a hybrid gene, i.e., a gene encoding a fusion protein. Specifically excluded from this definition are nucleic acids present in mixtures of (i) DNA molecules, (ii) transformed or transfected cells, and (iii) cell clones, e.g., as these occur in a DNA library such as a cDNA or genomic DNA library.

As used herein expression directed by a particular sequence is the transcription of an associated downstream sequence. If appropriate and desired for the associated sequence, there the term expression also encompasses translation (protein synthesis) of the transcribed RNA. When expression of a sequence of interest is "up-regulated," the expression is increased.

In the present context, a promoter is a DNA region which includes sequences sufficient to cause transcription of an associated (downstream) sequence. The promoter may be regulated, i.e., not constitutively acting to cause transcription of the associated sequence. If inducible, there are sequences present which mediate regulation of expression so that the associated sequence is transcribed only when an inducer molecule is present in the medium in or on which the organism is cultivated. In the present context, a transcription regulatory sequence includes a promoter sequence and the cis-active sequences necessary for regulated expression of an associated sequence in response to environmental signals.

One DNA portion or sequence is downstream of second DNA portion or sequence when it is located 3' of the second sequence. One DNA portion or sequence is upstream of a second DNA portion or sequence when it is located 5' of that sequence.

One DNA molecule or sequence and another are heterologous to another if the two are not derived from the same ultimate natural source. The sequences may be natural sequences, or at least one sequence can be designed by man, as in the case of a multiple cloning site region. The two sequences can be derived from two different species or one sequence can be produced by chemical synthesis provided that the nucleotide sequence of the synthesized portion was not derived from the same organism as the other sequence.

An isolated or substantially pure nucleic acid molecule or polynucleotide is a polynucleotide which is substantially separated from other polynucleotide sequences which naturally accompany a native *GALAT1* transcription regulatory sequence. The term embraces a polynucleotide sequence which has been removed from its naturally occurring environment, and includes recombinant or cloned DNA isolates, chemically synthesized analogues and analogues biologically synthesized by heterologous systems.

A polynucleotide is said to encode a polypeptide if, in its native state or when manipulated by methods known to those skilled in the art, it can be transcribed and/or translated to produce the polypeptide or a fragment thereof. The anti-sense strand of such a polynucleotide is also said to encode the sequence.

A nucleotide sequence is operably linked when it is placed into a functional relationship with another nucleotide sequence. For instance, a promoter is operably linked to a coding sequence if the promoter effects its transcription or expression. Generally, operably linked means that the sequences being linked are contiguous and, where necessary to join two protein coding regions, contiguous and in reading frame. However, it is well known that certain genetic elements, such as enhancers, may be operably linked even at a distance, i.e., even if not contiguous.

The term recombinant polynucleotide refers to a polynucleotide which is made by the combination of two otherwise separated segments of sequence accomplished by the artificial manipulation of isolated segments of polynucleotides by genetic engineering techniques or by chemical synthesis. In so doing one may join together polynucleotide segments of desired functions to generate a desired combination of functions.

Polynucleotide probes include an isolated polynucleotide attached to a label or reporter molecule and may be used to identify and isolate other *galat1* transcription regulatory sequences, for example, those from other species of *Aureobasidium* or other strains of *A. pullulans*. Probes comprising synthetic oligonucleotides or other polynucleotides may be derived from naturally occurring or recombinant single or double stranded nucleic acids or be chemically synthesized. Polynucleotide probes may be labeled by any of the methods known in the art, e.g., random hexamer labeling, nick translation, or the Klenow fill-in reaction.

Large amounts of the polynucleotides may be produced by replication in a suitable host cell. Natural or synthetic DNA fragments coding for a protein of interest are incorporated into recombinant polynucleotide constructs, typically DNA constructs, capable of introduction into

and replication in a prokaryotic or eukaryotic cell, Arabidopsis thaliana, wherein protein expression is desired. In addition to the Arabidopsis thaliana specifically exemplified herein, others can be used, including but not limited to, *A. pullulans* Y-117 [Li et al. (1997) *Fungal Genetics Newsletter* 44:29-32]; *A. pullulans* R106 [Thornewell et al. (1995) *Gene* 162:105-110]; *A. pullulans* NRRL Y-6220, Y-6754a, Y-12,974, YB-4026 and YB-4588 [Leathers et al. (1988) *J. Indus. Microbiol.* 3:231-239]; and *A. pullulans* CBS 58475 [Dobberstein and Emeis (1989) *Appl. Microbiol. Biotechnol.* 32:262-268]. Usually the construct is suitable for replication in a unicellular host, such as *A. pullulans* or a bacterium, but a multicellular eukaryotic host may also be appropriate, with or without integration within the genome of the host cell. Commonly used prokaryotic hosts include strains of *Escherichia coli*, although other prokaryotes, such as *Bacillus subtilis* or a pseudomonad, may also be used. Eukaryotic host cells include yeast, filamentous fungi, plant, insect, amphibian and avian species, but the regulated expression of a protein of interest a cell of the genus Aureobasidium, especially *A. pullulans*, is preferred. Such factors as ease of manipulation, ability to appropriately glycosylate expressed proteins, degree and control of protein expression, ease of purification of expressed proteins away from cellular contaminants or other factors influence the choice of the host cell.

The polynucleotides may also be produced by chemical synthesis, e.g., by the phosphoramidite method described by Beaucage and Caruthers (1981) *Tetra. Letts.*, 22: 1859-1862 or the triester method according to Matteuci et al. (1981) *J. Am. Chem. Soc.*, 103: 3185, and may be performed on commercial automated oligonucleotide synthesizers. A double-stranded fragment may be obtained from the single stranded product of chemical synthesis either by synthesizing the complementary strand and annealing the strand together under appropriate conditions or by adding the complementary strand using DNA polymerase with an appropriate primer sequence.

DNA constructs prepared for introduction into a prokaryotic or eukaryotic host will typically comprise a replication system (i.e. vector) recognized by the host, including the intended DNA fragment encoding the desired polypeptide, and will preferably also include transcription and translational initiation regulatory sequences operably linked to the polypeptide-encoding segment. Expression systems (expression vectors) may include, for example, an origin of replication or autonomously replicating sequence (ARS) and expression control sequences, a promoter, an enhancer and necessary processing information sites, such as ribosome-binding sites, RNA splice sites, polyadenylation sites, transcriptional terminator sequences, and mRNA stabilizing sequences. Signal peptides may also be included where appropriate from secreted polypeptides of the same or related species, which allow the protein to cross and/or lodge in cell membranes or be secreted from the cell.

An appropriate promoter and other necessary vector sequences will be selected so as to be functional in the host. Examples of workable combinations of cell lines and expression vectors are described in Sambrook et al. (1989) vide infra; Ausubel et al. (Eds.) (1995) *Current Protocols in Molecular Biology*, Greene Publishing and Wiley Interscience, New York; and Metzger et al. (1988) *Nature*, 334: 31-36. Many useful vectors for expression in bacteria, yeast, fungal, mammalian, insect, plant or other cells are well known in the art and may be obtained

such vendors as Stratagene, New England Biolabs, Promega Biotech, and others. In addition, the construct may be joined to an amplifiable gene (e.g., DHFR) so that multiple copies of the gene may be made. For appropriate enhancer and other expression control sequences, see also *Enhancers and Eukaryotic Gene Expression*, Cold Spring Harbor Press, N.Y. (1983). While such expression vectors may replicate autonomously, they may less preferably replicate by being inserted into the genome of the host cell.

Expression and cloning vectors will likely contain a selectable marker, that is, a gene encoding a protein necessary for the survival or growth of a host cell transformed with the vector. Although such a marker gene may be carried on another polynucleotide sequence co-introduced into the host cell, it is most often contained on the cloning vector. Only those host cells into which the marker gene has been introduced will survive and/or grow under selective conditions. Typical selection genes encode proteins that (a) confer resistance to antibiotics or other toxic substances, e.g., ampicillin, neomycin, methotrexate, etc.; (b) complement auxotrophic deficiencies; or (c) supply critical nutrients not available from complex media. The choice of the proper selectable marker will depend on the host cell; appropriate markers for different hosts are known in the art.

Recombinant host cells, in the present context, are those which have been genetically modified to contain an isolated DNA molecule of the instant invention. The DNA can be introduced by any means known to the art which is appropriate for the particular type of cell, including without limitation, transformation, lipofection or electroporation.

It is recognized by those skilled in the art that the DNA sequences may vary due to the degeneracy of the genetic code and codon usage. All DNA sequences which code for the GALAT1 signal peptide are included in this invention, including DNA sequences as given in herein having an ATG preceding the coding region for the mature protein, and including DNA sequences with and without the intron identified in the Specification.

Additionally, it will be recognized by those skilled in the art that allelic variations may occur in the DNA sequences which will not significantly change activity of the amino acid sequences of the peptides which the DNA sequences encode. All such equivalent DNA sequences are included within the scope of this invention and the definition of the regulated promoter region. The skilled artisan will understand that the sequence of the exemplified promoter sequence and the nucleotide sequence encoding the signal peptide can be used to identify and isolate additional, nonexemplified nucleotide sequences which are functionally equivalent to the sequences given in the Specification.

Hybridization procedures are useful for identifying polynucleotides with sufficient homology (i.e., homologs of GALAT1 in other plants) to the subject sequences to be useful as taught herein. The particular hybridization techniques is not essential to the subject invention. As improvements are made in hybridization techniques, they can be readily applied by one of ordinary skill in the art.

A probe and sample are combined in a hybridization buffer solution and held at an appropriate temperature until annealing occurs. Thereafter, the membrane is washed free of extraneous materials, leaving the sample and bound probe molecules typically detected and quantified by autoradiography and/or liquid scintillation counting. As is well known in the art, if the probe molecule and nucleic acid sample hybridize by forming a strong non-covalent bond between the two molecules, it can be reasonably assumed that the probe and sample are essentially identical, or completely complementary if the annealing and washing steps are carried out under conditions of high stringency. The probe's detectable label provides a means for determining whether hybridization has occurred.

In the use of the oligonucleotides or polynucleotides as probes, the particular probe is labeled with any suitable label known to those skilled in the art, including radioactive and non-radioactive labels. Typical radioactive labels include ^{32}P, ^{35}S, or the like. Non-radioactive labels include, for example, ligands such as biotin or thyroxine, as well as enzymes such as hydrolases or peroxidases, or a chemiluminescer such as luciferin, or fluorescent compounds like fluorescein and its derivatives. Alternatively, the probes can be made inherently fluorescent as described in International Application No. WO 93/16094.

Various degrees of stringency of hybridization can be employed. The more stringent the conditions, the greater the complementarity that is required for duplex formation. Stringency can be controlled by temperature, probe concentration, probe length, ionic strength, time, and the like. Preferably, hybridization is conducted under moderate to high stringency conditions by techniques well know in the art, as described, for example in Keller, G.H., M.M. Manak (1987) *DNA Probes*, Stockton Press, New York, NY., pp. 169-170, hereby incorporated by reference.

As used herein, moderate to high stringency conditions for hybridization are conditions which achieve the same, or about the same, degree of specificity of hybridization as the conditions employed by the current inventors. An example of high stringency conditions are hybridizing at 68° C in 5X SSC/5X Denhardt's solution/0.1% SDS, and washing in 0.2X SSC/0.1% SDS at room temperature. An example of conditions of moderate stringency are hybridizing at 68° C in 5X SSC/5X Denhardt's solution/0.1% SDS and washing at 42° C in 3X SSC. The parameters of temperature and salt concentration can be varied to achieve the desired level of sequence identity between probe and target nucleic acid. See, e.g., Sambrook et al. (1989) *vide infra* or Ausubel *et al.* (1995) *Current Protocols in Molecular Biology*, John Wiley & Sons, NY, NY, for further guidance on hybridization conditions.

Specifically, hybridization of immobilized DNA in Southern blots with ^{32}P-labeled gene specific probes was performed by standard methods (Maniatis *et al.*) In general, hybridization and subsequent washes were carried out under moderate to high stringency conditions that allowed for detection of target sequences with homology to the exemplified *GALAT1* sequences. For double-stranded DNA gene probes, hybridization can be carried out overnight at 20-25° C below the melting temperature (Tm) of the DNA hybrid in 6X SSPE 5X Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. The melting temperature is described by the following formula (Beltz, G.A., Jacobe, T.H., Rickbush, P.T., Chorbas, and F.C. Kafatos [1983] *Methods*

of Enzymology, R.Wu, L, Grossman and K Moldave [eds] Academic Press, New York 100:266-285).

Tm=81.5° C + 16.6 Log[Na+]+0.41(+G+C)-0.61(%formamide)-600/length of duplex in base pairs.

Washes are typically carried out as follows: twice at room temperature for 15 minutes in 1X SSPE, 0.1% SDS (low stringency wash), and once at TM-20° C for 15 minutes in 0.2X SSPE, 0.1% SDS (moderate stringency wash).

For oligonucleotide probes, hybridization was carried out overnight at 10-20° C below the melting temperature (Tm) of the hybrid 6X SSPE, 5X Denhardt's solution, 0.1% SDS, 0.1 mg/ml denatured DNA. Tm for oligonucleotide probes was determined by the following formula: TM(°C)=2(number T/A base pairs +4(number G/C base pairs) [Suggs, S.V., T. Miyake, E.H., Kawashime, M.J. Johnson, K. Itakura, and R.B. Wallace (1981) *ICB-UCLA Symp. Dev. Biol. Using Purified Genes*, D.D. Brown (ed.), Academic Press, New York, 23:683-693].

Washes were typically carried out as follows: twice at room temperature for 15 minutes 1X SSPE, 0.1% SDS (low stringency wash), and once at the hybridization temperature for 15 minutes in 1X SSPE, 0.1% SDS (moderate stringency wash).

In general, salt and/or temperature can be altered to change stringency. With a labeled DNA fragment >70 or so bases in length, the following conditions can be used: Low, 1 or 2X SSPE, room temperature; Low, 1 or 2X SSPE, 42° C; Moderate, 0.2X or 1X SSPE, 65° C; and High, 0.1X SSPE, 65° C.

Duplex formation and stability depend on substantial complementarity between the two strands of a hybrid, and, as noted above, a certain degree of mismatch can be tolerated. Therefore, the probe sequences of the subject invention include mutations (both single and multiple), deletions, insertions of the described sequences, and combinations thereof, wherein said mutations, insertions and deletions permit formation of stable hybrids with the target polynucleotide of interest. Mutations, insertions, and deletions can be produced in a given polynucleotide sequence in many ways, and those methods are known to an ordinarily skilled artisan. Other methods may become known in the future.

Thus, mutational, insertional, and deletional variants of the disclosed nucleotide sequences can be readily prepared by methods which are well known to those skilled in the art. These variants can be used in the same manner as the exemplified primer sequences so long as the variants have substantial sequence homology with the original sequence. As used herein, substantial sequence homology refers to homology which is sufficient to enable the variant polynucleotide to function in the same capacity as the polynucleotide from which the probe was derived. Preferably, this homology is greater than 80%, more preferably, this homology is greater than 85%, even more preferably this homology is greater than 90%, and most preferably, this homology is greater than 95%. The degree of homology or identity needed for the variant to

function in its intended capacity depends upon the intended use of the sequence. It is well within the skill of a person trained in this art to make mutational, insertional, and deletional mutations which are equivalent in function or are designed to improve the function of the sequence or otherwise provide a methodological advantage. Methods for confirming promoter activity an xylan- or xylose responsiveness are known in the art.

Polymerase Chain Reaction (PCR) is a repetitive, enzymatic, primed synthesis of a nucleic acid sequence. This procedure is well known and commonly used by those skilled in this art [see Mullis, U.S. Patent Nos. 4,683,195, 4,683,202, and 4,800,159; Saiki *et al.* (1985) *Science* 230:1350-1354]. PCR is based on the enzymatic amplification of a DNA fragment of interest that is flanked by two oligonucleotide primers that hybridize to opposite strands of the target sequence. The primers are oriented with the 3' ends pointing towards each other. Repeated cycles of heat denaturation of the template, annealing of the primers to their complementary sequences, and extension of the annealed primers with a DNA polymerase result in the amplification of the segment defined by the 5' ends of the PCR primers. Since the extension product of each primer can serve as a template for the other primer, each cycle essentially doubles the amount of DNA template produced in the previous cycle. This results in the exponential accumulation of the specific target fragment, up to several million-fold in a few hours. By using a thermostable DNA polymerase such as the *Taq* polymerase, which is isolated from the thermophilic bacterium *Thermus aquaticus*, the amplification process can be completely automated. Other enzymes which can be used are known to those skilled in the art.

It is well known in the art that the polynucleotide sequences of the present invention can be truncated and/or mutated such that certain of the resulting fragments and/or mutants of the original full-length sequence can retain the desired characteristics of the full-length sequence. A wide variety of restriction enzymes which are suitable for generating fragments from larger nucleic acid molecules are well known. In addition, it is well known that *Bal*31 exonuclease can be conveniently used for time-controlled limited digestion of DNA. See, for example, Maniatis (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, New York, pages 135-139, incorporated herein by reference. See also Wei et al. (1983 *J. Biol. Chem.* 258:13006-13512. By use of *Bal*31 exonuclease (commonly referred to as "erase-a-base" procedures), the ordinarily skilled artisan can remove nucleotides from either or both ends of the subject nucleic acids to generate a wide spectrum of fragments which are functionally equivalent to the subject nucleotide sequences. One of ordinary skill in the art can, in this manner, generate hundreds of fragments of controlled, varying lengths from locations all along the original *xynA* molecule. The ordinarily skilled artisan can routinely test or screen the generated fragments for their characteristics and determine the utility of the fragments as taught herein. It is also well known that the mutant sequences of the full length sequence, or fragments thereof, can be easily produced with site directed mutagenesis. See, for example, Larionov, O.A. and Nikiforov, V.G. (1982) *Genetika* 18(3):349-59; Shortle, D., DiMaio, D., and Nathans, D. (1981) *Annu. Rev. Genet.* 15:265-94; both incorporated herein by reference. The skilled artisan can routinely produce deletion-, insertion-, or substitution-type mutations and identify those resulting mutants which contain the desired characteristics of the full length wild-type sequence, or fragments thereof, i.e., those which retain promoter activity.

DNA sequences having at least 90, or at least 95% identity to the DNA sequences as given herein are most preferable. In addition, various functional equivalents are included in the definition of a GALAT1 coding sequences and transcription regulatory sequence. Following the teachings herein and using knowledge and techniques well known in the art, the skilled worker will be able to make a large number of operative embodiments having equivalent DNA sequences to those listed herein without the expense of undue experimentation.

As used herein percent sequence identity of two nucleic acids is determined using the algorithm of Karlin and Altschul (1990) Proc. Natl. Acad. Sci. USA 87:2264-2268, modified as in Karlin and Altschul (1993) Proc. Natl. Acad. Sci. USA 90:5873-5877. Such an algorithm is incorporated into the NBLAST and XBLAST programs of Altschul et al. (1990) J. Mol. Biol. 215:402-410. BLAST nucleotide searches are performed with the NBLAST program, score = 100, wordlength = 12, to obtain nucleotide sequences with the desired percent sequence identity. To obtain gapped alignments for comparison purposes, Gapped BLAST is used as described in Altschul et al. (1997) Nucl. Acids. Res. 25:3389-3402. When utilizing BLAST and Gapped

BLAST programs, the default parameters of the respective programs (NBLAST and XBLAST) are used. See, for example, the National Center for Biotechnology Information website on the internet.

Monoclonal or polyclonal antibodies, preferably monoclonal, specifically reacting with a protein of interest can be made by methods well known in the art. See, e.g., Harlow and Lane (1988) *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratories; Goding (1996) *Monoclonal Antibodies: Principles and Practice*, 3rd ed., Academic Press, San Diego, CA, and Ausubel et al. (1993) *Current Protocols in Molecular Biology*, Wiley Interscience/Greene Publishing, New York, NY.

Standard techniques for cloning, DNA isolation, amplification and purification, for enzymatic reactions involving DNA ligase, DNA polymerase, restriction endonucleases and the like, and various separation techniques are those known and commonly employed by those skilled in the art. A number of standard techniques are described in Sambrook et al. (1989) *Molecular Cloning*, Second Edition, Cold Spring Harbor Laboratory, Plainview, New York; Maniatis et al. (1982) *Molecular Cloning*, Cold Spring Harbor Laboratory, Plainview, New York; Wu (ed.) (1993) *Meth. Enzymol.* 218, Part I; Wu (ed.) (1979) *Meth. Enzymol.* 68; Wu et al. (eds.) (1983) *Meth. Enzymol.* 100 and 101; Grossman and Moldave (eds.) *Meth. Enzymol.* 65; Miller (ed.) (1972) *Experiments in Molecular Genetics*, Cold Spring Harbor Laboratory, Cold Spring Harbor, New York; Old and Primrose (1981) *Principles of Gene Manipulation*, University of California Press, Berkley; Schleif and Wensink (1982) *Practical Methods in Molecular Biology*; Glover (ed.) (1985) *DNA Cloning* Vol. I and II, IRL Press, Oxford, UK; Hames and Higgins (eds.) (1985) *Nucleic Acid Hybridization*, IRL Press, Oxford, UK; Setlow and Hollaender (1979) *Genetic Engineering: Principles and Methods*, Vols. 1-4, Plenum Press, New York; and Ausubel et al. (1992) *Current Protocols in Molecular Biology*, Greene/Wiley, New York, NY. Abbreviations and nomenclature, where employed, are deemed standard in the field and commonly used in professional journals such as those cited herein.

All references cited in the present application are incorporated by reference herein to the extent that there is no inconsistency with the present disclosure.

26 102
76

APPENDIX I

Sequence #1

Gene name: At3g61130
GeneBank accession # for reference: NM_115977.1
GI:18411855

Nucleotide sequence of Sequence #1:
Positions 314-2347.

ATTCCGA GGATGGCGCT AAAGCGAGGG CTATCTGGAG TTAACCGGAT
TAGAGGAAGT GGTGGTGGAT CTCGATCTGT GCTTGTGCTT CTCATATTTT TCTGTGTTTT
TGCACCTCTT TGCTTCTTTG TTGGCCGAGG AGTGTATATC GATTCCTCAA ATGATTATTC
AATTGTTTCT GTGAAGCAGA ATCTTGACTG GAGAGAACGT TTAGCAATGC AATCTGTTAG
ATCTCTTTTC TCGAAAGAGA TACTAGATGT TATAGCAACC AGCACAGCTG ATTTGGGTCC
TCTTAGCCTT GATTCTTTTA AGAAAAACAA TTTGTCTGCA TCATGGCGGG GAACCGGAGT
AGACCCCTCC TTTAGACATT CTGAGAATCC AGCAACTCCT GATGTCAAAT CTAATAACCT
GAATGAAAAA CGTGACAGCA TTTCAAAAGA TAGTATCCAT CAGAAAGTTG AGACACCTAC
AAAGATTCAC AGAAGGCAAC TAAGAGAGAA AAGGCGTGAG ATGCGGGCAA ATGAGTTAGT
TCAGCACAAT GATGACACGA TTTTGAAACT CGAAAATGCT GCCATTGAAC GCTCTAAGTC
TGTTGATTCT GCAGTCCTTG GTAAATACAG TATTTGGAGA AGAGAAAATG AGAATGACAA
CTCTGATTCA AATATACGCT TGATGCGGGA TCAAGTAATA ATGGCTAGAG TCTATAGTGG
GATTGCAAAA TTGAAAAACA AGAACGATTT GTTACAAGAA CTCCAGGCCC GACTTAAGGA
CAGCCAACGG GTTTTGGGGG AAGCAACATC TGATGCTGAT CTTCCTCGGA GTGCGCATGA
GAAACTCAGA GCCATGGGTC AAGTCTTGGC TAAAGCTAAG ATGCAGTTAT ATGACTGCAA
GCTGGTTACT GGAAAGCTGA GAGCAATGCT TCAGACTGCC GACGAACAAG TGAGGAGCTT
AAAGAAGCAG AGTACTTTTC TGGCTCAGTT AGCAGCAAAA ACCATTCCAA ATCCTATCCA
TTGCCTATCA ATGCGCTTGA CTATCGATTA CTATCTTCTG TCTCCGGAGA AAAGAAAATT
CCCTCGGAGT GAAAACCTAG AAAACCCTAA TCTTTATCAT TATGCCCTCT TTTCCGACAA
TGTATTAGCT GCATCAGTAG TTGTTAACTC AACCATCATG AATGCCAAGG ATCCTTCTAA
GCATGTTTTT CACCTTGTCA CGGATAAACT CAATTTCGGA GCAATGAACA TGTGGTTCCT
CCTAAACCCA CCCGGAAAGG CAACCATACA TGTGGAAAAC GTCGATGAGT TTAAGTGGCT
CAATTCATCT TACTGTCCTG TCCTTCGTCA GCTTGAATCT GCAGCAATGA GAGAGTACTA
TTTTAAAGCA GACCATCCAA CTTCAGGCTC TTCGAATCTA AAATACAGAA ACCCAAAGTA
TCTATCCATG TTGAATCACT TGAGATTCTA CCTCCCTGAG GTTTATCCCA AGCTGAACAA
AATCCTCTTC CTGGACGATG ACATCATTGT TCAGAAAGAC TTGACTCCAC TCTGGGAAGT
TAACCTGAAC GGCAAAGTCA ACGGTGCAGT CGAAACCTGT GGGGAAAGTT TCCACAGATT
CGACAAGTAT CTCAACTTTT CGAATCCTCA CATTGCGAGG AACTTCAATC CAAATGCTTG
TGGATGGGCT TATGGAATGA ACATGTTCGA CCTAAAGGAA TGGAAGAAGA GAGACATCAC
TGGTATATAC CACAAGTGGC AAAACATGAA TGAGAACAGG ACACTATGGA AGCTAGGGAC
ATTGCCACCA GGATTAATAA CATTCTACGG ATTAACACAT CCCTTAAACA AGGCGTGGCA
TGTGCTGGGA CTTGGATATA ACCCGAGTAT CGACAAGAAG GACATTGAGA ATGCAGCAGT
GGTTCACTAT AACGGGAACA TGAAACCATG GTTGGAGTTG GCAATGTCCA AATATCGGCC
GTATTGGACC AAGTACATCA AGTTTGATCA CCCATATCTT CGTCGTTGCA ACCTTCATGA
ATAAAAT

Amino Acid Sequence of Sequence #1:
GeneBank ID# NP_191672
Positions 1-673.

MALKRGLSGVNRIRGSGGGSRSVLVLLIFFCVFAPLCFFVGRGV
YIDSSNDYSIVSVKQNLDWRERLAMQSVRSLFSKEILDVIATSTADLGPLSLDSFKKN
NLSASWRGTGVDPSFRHSENPATPDVKSNNLNEKRDSISKDSIHQKVETPTKIHRRQL
REKRREMRANELVQHNDDTILKLENAAIERSKSVDSAVLGKYSIWRRENENDNSDSNI
RLMRDQVIMARVYSGIAKLKNKNDLLQELQARLKDSQRVLGEATSDADLPRSAHEKLR
AMGQVLAKAKMQLYDCKLVTGKLRAMLQTADEQVRSLKKQSTFLAQLAAKTIPNPIHC
LSMRLTIDYYLLSPEKRKFPRSENLENPNLYHYALFSDNVLAASVVVNSTIMNAKDPS
KHVFHLVTDKLNFGAMNMWFLLNPPGKATIHVENVDEFKWLNSSYCPVLRQLESAAMR
EYYFKADHPTSGSSNLKYRNPKYLSMLNHLRFYLPEVYPKLNKILFLDDDIIVQKDLT
PLWEVNLNGKVNGAVETCGESFHRFDKYLNFSNPHIARNFNPNACGWAYGMNMFDLKE
WKKRDITGIYHKWQNMNENRTLWKLGTLPPGLITFYGLTHPLNKAWHVLGLGYNPSID
KKDIENAAVVHYNGNMKPWLELAMSKYRPYWTKYIKFDHPYLRRCNLHE

Sequence #2

Gene name: At2g38650
GeneBank accession # for reference: AY050982
GI:15293096

Nucleotide sequence of Sequence #2:
Positions 198-2057.

ATG AAA GGC GGA GGC GGT GGT GGA GGA GGT GGT GGC GGA GGA AAA CGC CGG TGG
AAA GTT CTG GTG ATT GGA GTT TTG GTT CTT GTT ATT CTT TCT ATG CTT GTT CCT
CTT GCT TTC TTA CTC GGT CTT CAC AAT GGC TTT CAC TCT CCT GGA TTT GTC ACT
GTT CAA CCG GCT TCT TCA TTT GAG AGC TTT ACC AGA ATC AAT GCT ACT AAG CAT
ACA CAG AGA GAT GTA TCC GAA CGG GTC GAT GAG GTT CTT CAA AAA ATC AAT CCA
GTT CTT CCC AAG AAA AGC GAC ATA AAC GTG GGT TCC AGA GAT GTG AAT GCA ACA
AGC GGC ACT GAT TCT AAA AAA AGA GGA TTA CCA GTG TCC CCA ACT GTT GTT GCC
AAT CCA AGC CCT GCA AAT AAA ACA AAA TCG GAA GCC TCA TAT ACA GGT GTT CAG
AGG AAA ATA GTA AGT GGT GAT GAA ACT TGG AGA ACT TGT GAA GTG AAA TAT GGG
AGC TAC TGC CTC TGG AGG GAG GAA AAT AAG GAA CCA ATG AAA GAT GCC AAG GTG
AAG CAA ATG AAG GAC CAG CTG TTT GTG GCT AGA GCA TAC TAT CCC AGT ATT GCT
AAA ATG CCT TCT CAA AGC AAG TTG ACT CGG GAT ATG AAA CAG AAT ATC CAA GAG
TTT GAG CGT ATT CTT AGT GAA AGT TCT CAA GAT GCT GAC CTT CCA CCA CAG GTT
GAT AAA AAG TTG CAG AAG ATG GAA GCT GTA ATT GCA AAG GCA AAG TCT TTT CCA
GTC GAC TGT AAC AAT GTT GAC AAG AAA TTG AGA CAG ATC CTT GAT TTG ACT GAG
GAT GAA GCT AGT TTC CAC ATG AAA CAG AGT GTG TTC CTC TAC CAG CTT GCA GTA
CAG ACA ATG CCT AAG AGT CTT CAT TGC TTG TCA ATG CGA CTA ACT GTG GAA CAT
TTC AAG TCA GAT TCA CTT GAG GAT CCC ATT AGT GAG AAA TTT TCA GAT CCC TCA
TTA CTT CAC TTT GTT ATC ATC TCC GAT AAT ATA CTA GCA TCG TCC GTT GTG ATC
AAC TCA ACG GTT GTA CAT GCA AGG GAC AGT AAA AAC TTT GTT TTC CAT GTA CTG
ACA GAC GAG CAG AAT TAC TTT GCA ATG AAA CAA TGG TTT ATT AGG AAT CCT TGC
AAA CAA TCA ACT GTT CAA GTA TTG AAC ATT GAA AAA CTC GAG CTG GAC GAT TCT
GAT ATG AAA CTG TCT TTG TCT GCG GAG TTC CGT GTT TCC TTC CCC AGT GGT GAC
CTT TTG GCG TCT CAA CAG AAT AGA ACA CAC TAC TTA TCC CTT TTC TCT CAA TCT
CAC TAT CTT CTT CCC AAA TTA TTT GAC AAA TTG GAG AAG GTT GTG ATT CTG GAT
GAT GAC GTT GTA GTC CAG CGA GAC TTA TCT CCC CTT TGG GAC CTT GAT ATG GAA
GGG AAA GTG AAT GGC GCT GTT AAG TCG TGC ACT GTG AGA TTG GGT CAG CTA AGG
AGT CTC AAG AGA GGA AAT TTT GAT ACC AAT GCT TGT CTC TGG ATG TCT GGT TTG
AAT GTG GTT GAT CTT GCT AGA TGG AGG GCA TTG GGT GTT TCA GAA ACC TAT CAA
AAA TAT TAT AAA GAG ATG AGT AGT GGA GAT GAG TCG AGC GAA GCA ATT GCA TTG
CAG GCA AGC TTG CTC ACA TTT CAA GAC CAA GTA TAT GCT CTT GAC GAC AAA TGG
GCT CTA TCA GGG CTT GGT TAT GAC TAC TAC ATC AAT GCA CAA GCC ATA AAA AAC
GCA GCC ATA TTG CAC TAT AAC GGG AAC ATG AAG CCG TGG CTT GAG CTG GGA ATC
CCA AAT TAC AAA AAC TAT TGG AGA AGG CAT CTG AGT CGG GAA GAT CGG TTC TTG
AGT GAC TGT AAC GTG AAT CCT TGA

Amino Acid Sequence of Sequence #2:
GeneBank ID# NP_565893
Positions 1-619.

MKGGGGGGGGGGGGKRRWKVLVIGVLVLVILSMLVPLAFLLGLH
NGFHSPGFVTVQPASSFESFTRINATKHTQRDVSERVDEVLQKINPVLPKKSDINVGS
RDVNATSGTDSKKRGLPVSPTVVANPSPANKTKSEASYTGVQRKIVSGDETWRTCEVK
YGSYCLWREENKEPMKDAKVKQMKDQLFVARAYYPSIAKMPSQSKLTRDMKQNIQEFE
RILSESSQDADLPPQVDKKLQKMEAVIAKAKSFPVDCNNVDKKLRQILDLTEDEASFH
MKQSVFLYQLAVQTMPKSLHCLSMRLTVEHFKSDSLEDPISEKFSDPSLLHFVIISDN
ILASSVVINSTVVHARDSKNFVFHVLTDEQNYFAMKQWFIRNPCKQSTVQVLNIEKLE
LDDSDMKLSLSAEFRVSFPSGDLLASQQNRTHYLSLFSQSHYLLPKLFDKLEKVVILD
DDVVVQRDLSPLWDLDMEGKVNGAVKSCTVRLGQLRSLKRGNFDTNACLWMSGLNVVD
LARWRALGVSETYQKYYKEMSSGDESSEAIALQASLLTFQDQVYALDDKWALSGLGYD
YYINAQAIKNAAILHYNGNMKPWLELGIPNYKNYWRRHLSREDRFLSDCNVNP

Sequence #3

Gene name: At5g47780
GeneBank accession # for reference: AY056202
GI:15810326

Nucleotide sequence of Sequence #3:
Positions 66-1914.

ATG ATG GTG AAG CTT CGC AAT CTT GTT CTT TTC TTC ATG CTC CTC ACC GTC GTT
GCT CAT ATC CTT CTC TAC ACC GAT CCC GCT GCC TCC TTC AAG ACC CCC TTT TCT
AAA CGC GAT TTC CTC GAG GAC GTA ACC GCC TTG ACT TTC AAT TCC GAT GAG AAT
CGT TTG AAT CTT CTT CCT CGG GAA TCT CCC GCT GTG CTC AGA GGA GGA CTC GTC
GGT GCT GTC TAT TCC GAT AAG AAT TCA CGG CGG CTA GAC CAA TTG TCT GCT CGA
GTT CTT TCC GCC ACC GAC GAT GAT ACT CAC TCA CAT ACT GAC ATT TCC ATC AAA
CAA GTC ACT CAT GAT GCA GCC TCA GAC TCG CAT ATT AAT AGG GAA AAT ATG CAT
GTT CAA TTG ACC CAA CAA ACC TCT GAA AAA GTT GAT GAG CAA CCA GAG CCT AAT
GCT TTT GGA GCT AAG AAA GAT ACT GGA AAC GTG TTG ATG CCT GAT GCT CAA GTG
AGG CAT CTT AAA GAT CAG CTT ATT AGG GCA AAG GTT TAT CTT TCC CTT CCA TCT
GCA AAG GCC AAT GCT CAT TTT GTG AGA GAG CTT CGA CTC CGT ATT AAA GAA GTT
CAA CGG GCA CTT GCA GAT GCC TCC AAG GAT TCG GAT CTG CCA AAG ACT GCT ATA
GAA AAG CTA AAA GCA ATG GAG CAA ACA CTG GCC AAA GGC AAG CAG ATC CAA GAT
GAC TGT TCT ACA GTG GTC AAG AAG CTA CGT GCT ATG CTC CAC TCC GCA GAT GAG
CAG CTA CGG GTC CAT AAG AAG CAA ACC ATG TTT TTG ACT CAA TTG ACT GCT AAG
ACC ATT CCT AAA GGA CTT CAC TGC CTT CCT CTG CGC CTC ACT ACA GAC TAT TAT
GCT TTA AAT TCA TCT GAA CAA CAA TTT CCA AAT CAG GAG AAA CTA GAA GAT ACT
CAG CTG TAT CAC TAT GCC CTT TTC TCT GAT AAT GTT TTG GCT ACG TCA GTT GTT
GTT AAC TCT ACC ATA ACC AAT GCA AAG CAT CCC TTA AAG CAT GTC TTC CAC ATC
GTC ACA GAC AGA CTC AAT TAT GCG GCA ATG AGG ATG TGG TTC CTG GAC AAT CCA
CCT GGC AAA GCC ACC ATC CAG GTT CAG AAT GTT GAA GAA TTT ACA TGG CTG AAT
TCA AGC TAC AGT CCC GTT CTC AAA CAG CTT AGT TCT AGA TCG ATG ATA GAT TAT
TAC TTC AGA GCC CAC CAT ACA AAT TCA GAC ACC AAC TTG AAG TTC CGG AAT CCA
AAA TAC TTA TCG ATC CTT AAT CAT CTT CGT TTT TAC TTG CCT GAG ATC TTT CCC
AAG CTC AGC AAA GTG CTC TTC TTG GAT GAT GAT ATA GTT GTG CAG AAG GAC CTT
TCT GGT CTT TGG TCA GTT GAT CTG AAA GGT AAT GTT AAC GGT GCT GTA GAG ACG
TGT GGG GAA AGC TTT CAT CGC TTT GAC CGT TAT CTG AAC TTC TCA AAT CCA CTC
ATT TCC AAG AAC TTT GAC CCT CGA GCT TGT GGT TGG GCG TAT GGT ATG AAT GTC
TTT GAT CTG GAT GAA TGG AAG AGG CAA AAC ATC ACA GAA GTT TAT CAT CGA TGG
CAG GAT CTG AAT CAA GAC CGA GAA TTG TGG AAG CTA GGG ACG TTG CCG CCT GGT
CTA ATC ACA TTT TGG AGA CGA ACA TAT CCG CTA GAC CGG AAA TGG CAC ATA CTA
GGG CTT GGA TAC AAC CCG AGT GTG AAC CAA AGG GAT ATT GAG AGG GCA GCC GTG
ATA CAC TAT AAT GGC AAC CTC AAA CCA TGG CTA GAG ATT GGG ATT CCA AGA TAC
AGA GGC TTC TGG TCA AAG CAT GTA GAC TAT GAG CAC GTT TAT CTC AGA GAA TGC
AAC ATC AAT CCT TAG

Amino Acid Sequence of Sequence #3:
Genebank ID# AAL07051
Positions 1-616.

MMVKLRNLVLFFMLLTVVAHILLYTDPAASFKTPFSKRDFLEDV
TALTFNSDENRLNLLPRESPAVLRGGLVGAVYSDKNSRRLDQLSARVLSATDDDTHSH
TDISIKQVTHDAASDSHINRENMHVQLTQQTSEKVDEQPEPNAFGAKKDTGNVLMPDA
QVRHLKDQLIRAKVYLSLPSAKANAHFVRELRLRIKEVQRALADASKDSDLPKTAIEK
LKAMEQTLAKGKQIQDDCSTVVKKLRAMLHSADEQLRVHKKQTMFLTQLTAKTIPKGL
HCLPLRLTTDYYALNSSEQQFPNQEKLEDTQLYHYALFSDNVLATSVVVNSTITNAKH
PLKHVFHIVTDRLNYAAMRMWFLDNPPGKATIQVQNVEEFTWLNSSYSPVLKQLSSRS
MIDYYFRAHHTNSDTNLKFRNPKYLSILNHLRFYLPEIFPKLSKVLFLDDDIVVQKDL
SGLWSVDLKGNVNGAVETCGESFHRFDRYLNFSNPLISKNFDPRACGWAYGMNVFDLD
EWKRQNITEVYHRWQDLNQDRELWKLGTLPPGLITFWRRTYPLDRKWHILGLGYNPSV
NQRDIERAAVIHYNGNLKPWLEIGIPRYRGFWSKHVDYEHVYLRECNINP

Sequence #4

Gene name: At1g06780
GeneBank accession # for reference: AY091452
GI:20259310

Nucleotide sequence of Sequence #4:
Positions 1-1770.

ATG AAA CAA ATT CGT CGA TGG CAG AGG ATT TTG ATC CTC GCT CTG CTA TCG ATA
TCA GTA TTC GCT CCG CTT ATT TTC GTA TCG AAT CGG CTT AAG AGC ATC ACT CCC
GTT GGT CGT AGA GAA TTT ATT GAA GAG TTA TCC AAA ATT AGA TTC ACG ACA AAT
GAC CTT AGA CTT AGC GCT ATT GAA CAT GAG GAT GGA GAA GGC TTG AAG GGG CCA
AGG CTC ATT CTC TTC AAG GAT GGG GAG TTT AAT TCG TCT GCT GAA AGT GAT GGT
GGT AAT ACT TAC AAA AAC AGG GAA GAA CAA GTG ATT GTT TCA CAG AAG ATG ACA
GTT AGC TCT GAT GAA AAG GGT CAA ATT CTA CCA ACA GTC AAC CAA CTT GCT AAT
AAA ACG GAT TTC AAG CCC CCT TTA TCT AAG GGT GAA AAG AAC ACA AGG GTT CAG
CCC GAC AGA GCA ACA GAT GTG AAA ACG AAG GAG ATC AGA GAC AAA ATT ATT CAA
GCT AAA GCC TAC CTG AAT TTC GCT CCA CCT GGA AGT AAC TCT CAA GTT GTG AAG
GAG TTG AGA GGT CGG CTG AAA GAG CTG GAA CGG TCT GTT GGT GAT GCA ACA AAG
GAC AAG GAC TTA TCA AAG GGC GCT CTC CGC AGG GTG AAG CCC ATG GAA AAT GTG
TTA TAT AAG GCT AGT CGT GTC TTT AAC AAT TGC CCT GCC ATC GCT ACC AAA CTC
CGT GCC ATG AAT TAT AAC ACA GAA GAA CAA GTT CAG GCG CAG AAA AAT CAA GCA
GCG TAT CTA ATG CAG CTT GCA GCA AGG ACC ACC CCA AAA GGG CTT CAC TGT CTC
TCA ATG CGG CTG ACA TCA GAA TAC TTT TCA CTG GAT CCT GAA AAA AGG CAG ATG
CCT AAC CAG CAA AAT TAT TTT GAC GCT AAT TTC AAT CAT TAT GTT GTC TTC TCT
GAC AAT GTT TTG GCT TCT TCA GTC GTT GTT AAC TCT ACG ATA TCT TCA TCA AAG
GAG CCA GAA AGA ATA GTC TTC CAT GTC GTG ACT GAT TCA CTT AAT TAC CCA GCA
ATC TCA ATG TGG TTT CTG CTA AAC ATT CAA AGT AAA GCT ACT ATC CAA ATC CTA
AAC ATT GAT GAT ATG GAT GTC CTG CCT AGA GAT TAT GAT CAA TTA CTG ATG AAG
CAA AAC TCT AAT GAC CCA AGA TTC ATT TCT ACA CTC AAT CAC GCA CGC TTC TAT
CTC CCG GAT ATA TTC CCG GGT TTG AAC AAG ATG GTA CTC TTG GAC CAT GAT GTA
GTT GTT CAA AGA GAT TTA AGT AGA CTG TGG AGC ATT GAT ATG AAA GGA AAG GTG
GTT GGA GCT GTA GAG ACT TGT CTT GAA GGT GAA TCT TCA TTT CGA TCA ATG AGC
ACA TTT ATT AAT TTC TCA GAC ACA TGG GTC GCT GGG AAA TTT AGT CCT AGA GCT
TGC ACA TGG GCT TTC GGG ATG AAT CTA ATT GAT CTC GAA GAA TGG AGA ATA CGG
AAG TTG ACT TCT ACA TAC ATA AAA TAC TTC AAC CTG GGA ACA AAG AGA CCA TTG
TGG AAA GCT GGG AGC TTA CCA ATA GGT TGG TTG ACT TTC TAT AGG CAA ACA TTA
GCA TTG GAC AAG AGA TGG CAT GTG ATG GGG TTA GGT CGC GAA TCA GGA GTC AAA
GCG GTT GAC ATC GAA CAA GCG GCA GTT ATA CAC TAC GAT GGG GTC ATG AAG CCG
TGG TTG GAC ATT GGA AAA GAG AAT TAC AAA CGT TAC TGG AAC ATA CAC GTC CCT
TAC CAT CAC ACC TAC TTG CAA CAG TGC AAT CTT CAA GCT TGA

Amino Acid Sequence of Sequence #4:
Genebank ID# NP_563771
Positions 1-589.

MKQIRRWQRILILALLSISVFAPLIFVSNRLKSITPVGRREFIE
ELSKIRFTTNDLRLSAIEHEDGEGLKGPRLILFKDGEFNSSAESDGGNTYKNREEQVI
VSQKMTVSSDEKGQILPTVNQLANKTDFKPPLSKGEKNTRVQPDRATDVKTKEIRDKI
IQAKAYLNFAPPGSNSQVVKELRGRLKELERSVGDATKDKDLSKGALRRVKPMENVLY
KASRVFNNCPAIATKLRAMNYNTEEQVQAQKNQAAYLMQLAARTTPKGLHCLSMRLTS
EYFSLDPEKRQMPNQQNYFDANFNHYVVFSDNVLASSVVVNSTISSSKEPERIVFHVV
TDSLNYPAISMWFLLNIQSKATIQILNIDDMDVLPRDYDQLLMKQNSNDPRFISTLNH
ARFYLPDIFPGLNKMVLLDHDVVVQRDLSRLWSIDMKGKVVGAVETCLEGESSFRSMS
TFINFSDTWVAGKFSPRACTWAFGMNLIDLEEWRIRKLTSTYIKYFNLGTKRPLWKAG
SLPIGWLTFYRQTLALDKRWHVMGLGRESGVKAVDIEQAAVIHYDGVMKPWLDIGKEN
YKRYWNIHVPYHHTYLQQCNLQA

Sequence #5

Gene name: At1g18580
GeneBank accession # for reference: AY062444
GI:17064735

Nucleotide sequence of Sequence #5:
Positions 281-1892.

ATG AGG CGG TGG CCG GTG GAT CAC CGG CGG CGA GGT AGA AGG AGA TTG TCG AGT
TGG ATA TGG TTT CTC CTT GGT TCT TTC TCT GTC GCT GGT TTA GTT CTC TTC ATC
GTT CAG CAT TAT CAC CAT CAA CAA GAT CCA TCC CAG CTT TTA CTT GAG AGA GAC
ACG AGA ACC GAA ATG GTA TCT CCT CCC CAT TTA AAC TTC ACG GAA GAG GTC ACA
AGT GCT TCC TCC TTC TCT AGG CAG TTA GCA GAG CAA ATG ACA CTT GCC AAA GCT
TAT GTG TTT ATA GCT AAA GAG CAT AAT AAT CTT CAT TTA GCT TGG GAA TTG AGT
TCT AAG ATC AGA AGT TGT CAG CTT TTG CTT TCC AAA GCA GCT ATG AGA GGA CAA
CCT ATT TCG TTT GAT GAG GCT AAA CCG ATT ATT ACT GGT CTA TCA GCT CTT ATC
TAC AAG GCT CAA GAT GCA CAT TAT GAT ATT GCC ACC ACT ATG ATG ACC ATG AAA
TCT CAC ATC CAA GCA CTT GAA GAG CGT GCA AAT GCA GCT ACT GTT CAG ACC ACA
ATA TTT GGG CAA TTG GTT GCT GAG GCA TTA CCA AAG AGC CTC CAC TGT TTG ACG
ATA AAG CTC ACA TCT GAT TGG GTA ACA GAG CCA TCT CGC CAT GAA CTG GCA GAT
GAG AAC AGA AAC TCA CCT AGA CTT GTC GAC AAC AAC CTC TAC CAC TTC TGC ATC
TTC TCG GAC AAC GTG ATT GCC ACC TCG GTT GTT GTT AAT TCA ACT GTC TCG AAT
GCT GAT CAT CCA AAG CAG CTT GTT TTC CAC ATA GTG ACG AAT CGA GTG AGC TAC
AAA GCT ATG CAG GCC TGG TTT CTA AGT AAT GAC TTC AAG GGC TCA GCA ATA GAG
ATC AGG AGC GTA GAG GAG TTT TCT TGG TTG AAT GCT TCA TAT TCT CCT GTT GTT
AAG CAA CTG CTG GAC ACA GAT GCA AGA GCT TAC TAT TTC GGG GAA CAG ACA AGT
CAA GAT ACG ATT TCC GAG CCA AAA GTG AGG AAC CCA AAG TAC TTG TCA TTA CTG
AAC CAT CTC AGA TTC TAC ATT CCG GAG ATC TAT CCA CAG CTA GAG AAG ATT GTT
TTC CTA GAC GAT GAT GTT GTT GTT CAG AAA GAT TTG ACT CCA CTC TTC TCC TTG
GAT CTG CAT GGA AAC GTC AAT GGA GCT GTG GAA ACA TGT CTT GAA GCC TTT CAC
CGA TAT TAC AAG TAT CTA AAT TTC TCG AAC CCA CTC ATC AGC TCA AAG TTC GAC
CCA CAA GCA TGT GGA TGG GCT TTT GGT ATG AAC GTT TTT GAT CTG ATC GCT TGG
AGG AAT GCA AAC GTG ACT GCT CGG TAC CAT TAC TGG CAA GAT CAG AAC AGA GAA
CGA ACG CTT TGG AAA CTC GGG ACA CTC CCT CCA GGT CTA CTA TCT TTC TAT GGT
CTC ACA GAG CCA CTG GAC AGA AGA TGG CAT GTC TTG GGT TTA GGT TAC GAT GTG
AAC ATC GAT AAC CGT CTG ATC GAA ACA GCA GCT GTG ATT CAC TAT AAT GGT AAC
ATG AAG CCT TGG CTA AAG CTG GCT ATT GGT AGG TAT AAA CCT TTC TGG TTA AAG
TTT TTG AAC TCG AGC CAT CCT TAT TTA CAA GAT TGT GTC ACA GCT TAA

Amino Acid Sequence of Sequence #4:
Genebank ID# aak93644.
Positions 1-537.

MRRWPVDHRRRGRRRLSSWIWFLLGSFSVAGLVLFIVQHYHHQQDPSQLLLERDTRTEMVSPPHLNFTEEV
TSASSFSRQLAEQMTLAKAYVFIAKEHNNLHLAWELSSKIRSCQLLLSKAAMRGQPISFDEAKPIITGLSA
LIYKAQDAHYDIATTMMTMKSHIQALEERANAATVQTTIFGQLVAEALPKSLHCLTIKLTSDWVTEPSRHE
LADENRNSPRLVDNNLYHFCIFSDNVIATSVVVNSTVSNADHPKQLVFHIVTNRVSYKAMQAWFLSNDFKG
SAIEIRSVEEFSWLNASYSPVVKQLLDTDARAYYFGEQTSQDTISEPKVRNPKYLSLLNHLRFYIPEIYPQ
LEKIVFLDDDVVVQKDLTPLFSLDLHGNVNGAVETCLEAFHRYYKYLNFSNPLISSKFDPQACGWAFGMNV
FDLIAWRNANVTARYHYWQDQNRERTLWKLGTLPPGLLSFYGLTEPLDRRWHVLGLGYDVNIDNRLIETAA
VIHYNGNMKPWLKLAIGRYKPFWLKFLNSSHPYLQDCVTA

Sequence #6

Gene name: At2g20810
GeneBank accession # for reference: AY059085.1 GI:16323393
Nucleotide sequence of Sequence #6:
Positions 1-1609.

ATG AGA AGG AGA GGA GGG GAT AGT TTC CGG AGA GCT GGA CGG AGG AAG ATC TCG
AAT GTG GTA TGG TGG GTT CTC TCT GGT ATT GCC CTC CTG CTC TTC TTT CTC ATT
CTC TCC AAA GCT GGT CAT ATT GAA CCT AGA CCC TCT ATT CCT AAG CGA CGT TAC
CGT AAT GAC AAA TTT GTA GAG GGT ATG AAT ATG ACT GAG GAA ATG TTG AGT CCT
ACT TCC GTT GCT CGT CAA GTT AAT GAT CAG ATT GCT CTT GCT AAA GCT TTT GTT
GTC ATT GCT AAA GAA AGT AAG AAT CTT CAG TTT GCT TGG GAC TTA AGT GCT CAG
ATC CGT AAC TCT CAG TTG CTT TTA TCG AGT GCT GCT ACT AGG AGA AGT CCC TTG
ACT GTC TTG GAA TCT GAG TCT ACT ATT CGT GAC ATG GCT GTT TTG TTA TAT CAA
GCT CAG CAG CTT CAC TAT GAT AGT GCT ACT ATG ATT ATG AGG CTT AAG GCC TCG
ATT CAG GCT CTT GAA GAA CAA ATG AGT TCC GTT AGC GAG AAG AGT TCC AAG TAT
GGA CAG ATT GCT GCT GAG GAA GTG CCT AAG AGT CTT TAC TGT CTT GGT GTT CGT
CTC ACT ACC GAA TGG TTT CAG AAT TTA GAC TTA CAG AGA ACT CTT AAG GAA AGG
AGT CGT GTT GAT TCG AAA CTC ACG GAT AAC AGT CTC TAC CAT TTC TGT GTG TTT
TCC GAT AAC ATT ATT GCT ACT TCT GTT GTG GTT AAT TCT ACT GCT CTC AAT TCC
AAG GCC CCT GAG AAA GTT GTG TTT CAT CTT GTG ACT AAT GAG ATC AAC TAT GCT
GCA ATG AAG GCT TGG TTC GCC ATT AAT ATG GAC AAC CTC AGA GGA GTC ACT GTG
GAG GTT CAG AAG TTC GAG GAT TTC TCA TGG CTG AAT GCT TCC TAT GTT CCG GTC
CTC AAG CAG CTG CAA GAC TCT GAT ACG CAA AGC TAT TAT TTC TCT GGA CAC AAC
GAT GAT GGG CGC ACT CCA ATC AAA TTC AGG AAC CCC AAG TAT CTT TCC ATG CTC
AAC CAT CTT AGG TTC TAC ATC CCT GAA GTG TTT CCT GCG CTG AAG AAG GTG GTC
TTT CTT GAT GAT GAT GTT GTA GTT CAG AAG GAT CTT TCA TCT CTC TTT TCG ATC
GAT TTA AAC AAA AAT GTG AAC GGG GCT GTT GAG ACC TGC ATG GAG ACC TTC CAC
CGC TAC CAC AAG TAC TTG AAC TAT TCT CAT CCT CTC ATA CGC TCC CAC TTT GAT
CCA GAT GCG TGT GGG TGG GCG TTT GGA ATG AAC GTC TTT GAT TTA GTT GAG TGG
AGG AAG AGA AAT GTG ACC GGC ATA TAC CAC TAC TGG CAA GAA AAA AAC GTG GAC
CGG ACC TTA TGG AAA CTG GGA ACA CTA CCT CCA GGA CTT CTG ACA TTT TAC GGG
TTA ACA GAG GCA CTA GAG GCG TCC TGG CAT ATC CTG GGA TTG GGA TAC ACG AAT
GTG GAT GCT CGT GTG ATA GAG AAA GGA GCT GTT CTT CAC TTC AAT GGG AAC TTA
AAG CCA TGG TTG AAG ATC GGG ATA GAG AAG TAC AAA CCT TTG TGG GAG AGA TAC
GTT GAT TAC ACT TCT CCT TTT ATG CAA CAA TGC AAT TTT CAT

Amino Acid Sequence of Sequence #6:
Genebank ID# NP_565485
Positions 1-536.

MRRRGGDSFRRAGRRKISNVVWWVLSGIALLLFFLILSKAGHIEPRPSIPKRRYRNDKFVEGMNMTEEMLS
PTSVARQVNDQIALAKAFVVIAKESKNLQFAWDLSAQIRNSQLLLSSAATRRSPLTVLESESTIRDMAVLL
YQAQQLHYDSATMIMRLKASIQALEEQMSSVSEKSSKYGQIAAEEVPKSLYCLGVRLTTEWFQNLDLQRTL
KERSRVDSKLTDNSLYHFCVFSDNIIATSVVVNSTALNSKAPEKVVFHLVTNEINYAAMKAWFAINMDNLR
GVTVEVQKFEDFSWLNASYVPVLKQLQDSDTQSYYFSGHNDDGRTPIKFRNPKYLSMLNHLRFYIPEVFPA
LKKVVFLDDDVVVQKDLSSLFSIDLNKNVNGAVETCMETFHRYHKYLNYSHPLIRSHFDPDACGWAFGMNV
FDLVEWRKRNVTGIYHYWQEKNVDRTLWKLGTLPPGLLTFYGLTEALEASWHILGLGYTNVDARVIEKGAV
LHFNGNLKPWLKIGIEKYKPLWERYVDYTSPFMQQCNFH

Sequence #7

Gene name: At2g30575
GeneBank accession # for reference: NM_147355.1 GI:22326004

Nucleotide sequence of Sequence #7:
Positions 303-2135.

ATG AAT CAA GTT CGT CGT TGG CAG AGG ATT CTG ATC CTC TCG CTG CTA TTG TTA
TCT GTT TTA GCT CCG ATT GTT TTC GTT TCG AAT CGG CTC AAG AGC ATC ACT TCC
GTC GAT AGA GGA GAA TTC ATT GAA GAA TTA TCC GAC ATT ACA GAT AAG ACC GAG
GAT GAA CTT AGA CTT ACT GCT ATT GAA CAG GAC GAA GAA GGC TTG AAG GAG CCT
AAA CGT ATT CTG CAG GAT CGA GAT TTT AAT TCT GTG GTT TTG TCA AAT TCC TCT
GAT AAA AGT AAT GAT ACT GTG CAG TCT AAT GAG GGA GAC CAA AAA AAC TTT CTC
TCA GAA GTT GAT AAG GGA AAT AAT CAC AAA CCA AAG GAG GAA CAA GCA GTT TCA
CAG AAA ACC ACA GTA AGC TCG AAT GCG GAG GTG AAA ATT TCA GCA AGA GAT ATT
CAA CTT AAT CAT AAA ACG GAA TTC CGA CCC CCT TCA AGT AAG AGT GAA AAG AAT
ACA AGG GTT CAA CTT GAA AGA GCA ACA GAT GAG AGG GTA AAG GAG ATC AGA GAC
AAA ATT ATC CAA GCG AAA GCC TAT CTG AAT TTG GCC CTA CCT GGG AAT AAC TCC
CAA ATC GTA AAG GAG TTG AGA GTT CGA ACG AAA GAG CTG GAA CGG GCT ACT GGT
GAT ACT ACC AAG GAT AAA TAT TTG CCA AAG AGC TCT CCT AAC AGA TTG AAG GCC
ATG GAA GTT GCG TTA TAC AAG GTC AGC CGT GCC TTT CAC AAC TGC CCT GCC ATT
GCT ACC AAA CTC CAA GCC ATG ACT TAT AAA ACC GAA GAA CAA GCT CGG GCG CAG
AAG AAA CAA GCA GCA TAT TTA ATG CAG CTT GCA GCA AGG ACT ACC CCA AAA GGG
CTT CAT TGT CTC TCA ATG CGG TTG ACA ACA GAA TAT TTT ACC CTG GAT CAC GAA
AAA AGG CAG CTT TTG CAA CAA AGT TAT AAT GAT CCT GAT CTC TAC CAT TAC GTA
GTC TTC TCT GAC AAT GTT TTG GCC TCT TCG GTT GTT GTT AAC TCT ACA ATC TCC
TCA TCA AAG GAA CCG GAT AAA ATA GTA TTC CAT GTG GTG ACA GAT TCA CTC AAT
TAC CCA GCA ATC TCA ATG TGG TTT TTA CTA AAC CCA AGT GGC AGA GCT TCA ATC
CAA ATC CTA AAC ATT GAT GAA ATG AAT GTC CTG CCA TTG TAC CAT GCT GAA TTG
CTG ATG AAG CAA AAT TCA AGT GAC CCA AGA ATC ATT TCA GCG CTC AAC CAT GCA
CGC TTC TAT CTC CCA GAT ATC TTC CCA GGT CTA AAC AAG ATC GTA CTC TTC GAT
CAT GAT GTA GTA GTG CAA AGG GAT CTA ACT AGA CTG TGG AGC CTT GAT ATG ACG
GGG AAA GTT GTT GGA GCT GTA GAG ACT TGT CTT GAA GGT GAT CCT TCA TAT CGT
TCG ATG GAC TCA TTC ATT AAT TTC TCA GAT GCA TGG GTT TCT CAG AAA TTT GAT
CCC AAG GCT TGC ACT TGG GCA TTC GGG ATG AAT CTA TTT GAT CTC GAA GAA TGG
AGA AGA CAG GAG TTG ACT TCT GTA TAC CTG AAA TAC TTC GAC CTG GGA GTA AAA
GGA CAT CTG TGG AAA GCA GGG GGA TTG CCA GTA GGT TGG TTG ACT TTT TTC GGG
CAA ACG TTT CCG TTG GAA AAG AGA TGG AAC GTG GGT GGG TTA GGT CAC GAA TCA
GGA CTC AGG GCA AGC GAC ATC GAA CAA GCA GCG GTT ATA CAC TAC GAC GGG ATC
ATG AAA CCA TGG CTG GAC ATC GGT ATA GAC AAG TAC AAG CGC TAC TGG AAC ATA
CAT GTA CCT TAC CAT CAC CCT CAC TTA CAA CGG TGC AAC ATT CAC GAT TGA

Amino Acid Sequence of Sequence #7:
Genebank ID# NP_671901
Positions 1-610.

MNQVRRWQRILILSLLLLSVLAPIVFVSNRLKSITSVDRGEFIEELSDITDKTEDELRLTAIEQDEEGLKE
PKRILQDRDFNSVVLSNSSDKSNDTVQSNEGDQKNFLSEVDKGNNHKPKEEQAVSQKTTVSSNAEVKISAR
DIQLNHKTEFRPPSSKSEKNTRVQLERATDERVKEIRDKIIQAKAYLNLALPGNNSQIVKELRVRTKELER
ATGDTTKDKYLPKSSPNRLKAMEVALYKVSRAFHNCPAIATKLQAMTYKTEEQARAQKKQAAYLMQLAART
TPKGLHCLSMRLTTEYFTLDHEKRQLLQQSYNDPDLYHYVVFSDNVLASSVVVNSTISSSKEPDKIVFHVV
TDSLNYPAISMWFLLNPSGRASIQILNIDEMNVLPLYHAELLMKQNSSDPRIISALNHARFYLPDIFPGLN
KIVLFDHDVVVQRDLTRLWSLDMTGKVVGAVETCLEGDPSYRSMDSFINFSDAWVSQKFDPKACTWAFGMN
LFDLEEWRRQELTSVYLKYFDLGVKGHLWKAGGLPVGWLTFFGQTFPLEKRWNVGGLGHESGLRASDIEQA
AVIHYDGIMKPWLDIGIDKYKRYWNIHVPYHHPHLQRCNIHD.

Sequence #8

Gene name: At2g46480
GeneBank accession # for reference: NM_130212.2 GI:22326493

Nucleotide sequence of Sequence #8:
Positions 1-1587.

ATG ACT GAT GCT TGT TGT TTG AAG GGA AAC GAG GAC AAA ATG GTT CCT CGT TTT
GGT CAT GGA ACC TGG ATA GGA AAA GCA TTT AAT GAT ACA CCA GAG ATG TTG CAT
GAA AGG AGT CTG AGA CAG GAA AAA AGA TTG GAA AGG GCT AAT GAG CTG ATG AAT
GAT GAT AGT CTG CAA AAG CTT GAG ACG GCA GCC ATG GCA CGT TCC AGA TCT GTC
GAT TCT GCA CCA CTA GGA AAC TAC ACC ATT TGG AAA AAT GAA TAC CGG AGG GGC
AAG AGT TTT GAA GAT ATG TTA CGT TTG ATG CAA GAT CAA ATC ATC ATG GCA CGA
GTT TAC AGT GGA CTT GCA AAG TTT ACA AAC AAT CTC GCC TTG CAC CAA GAG ATA
GAA ACA CAA CTA ATG AAA CTA GCT TGG GAG GAA GAA TCT ACT GAT ATT GAT CAG
GAG CAG AGA GTA CTT GAC AGT ATA AGA GAC ATG GGA CAA ATA CTG GCT AGA GCA
CAC GAG CAG CTA TAT GAA TGC AAG TTG GTG ACA AAT AAG TTG AGA GCA ATG CTA
CAA ACA GTT GAA GAT GAA CTC GAA AAC GAG CAG ACT TAT ATA ACG TTC TTG AGT
CAG CTA GCT TCC AAG GCA CTA CCA GAT GCT ATC CAC TGC TTG ACC ATG CGC TTG
AAT CTA GAG TAT CAT CTC CTG CCT TTA CCG ATG AGA AAT TTT CCA AGG AGG GAG
AAT TTG GAG AAT CCA AAA CTT TAC CAC TAC GCT CTC TTC TCT GAT AAT GTA CTG
GCT GCA TCA GTT GTT GTC AAC TCC ACA GTC ATG AAT GCA CAG GAT CCT TCA AGG
CAT GTT TTC CAC CTT GTG ACT GAT AAG CTC AAC TTT GGA GCA ATG AGT ATG TGG
TTT CTG TTG AAC CCT CCT GGA GAA GCG ACC ATC CAT GTC CAA AGG TTT GAA GAT
TTT ACT TGG CTC AAC TCA TCT TAC TCT CCA GTT TTG AGT CAG CTC GAG TCA GCA
GCT ATG AAG AAG TTC TAC TTC AAG ACA GCG AGG TCT GAA TCA GTT GAA TCA GGC
TCA GAA AAC CTC AAG TAC CGG TAC CCG AAA TAC ATG TCA ATG CTT AAC CAC CTG
AGG TTC TAC ATC CCT AGG ATC TTC CCA AAG TTG GAG AAA ATC TTG TTT GTT GAC
GAT GAT GTG GTT GTT CAG AAG GAT TTA ACT CCC CTA TGG TCC ATT GAT CTT AAA
GGG AAA GTG AAT GAA AAC TTT GAT CCC AAG TTC TGC GGA TGG GCT TAT GGG ATG
AAC ATC TTC GAC CTG AAA GAA TGG AAG AAG AAC AAC ATT ACA GAA ACT TAT CAC
TTT TGG CAA AAC CTG AAC GAA AAC CGG ACT CTA TGG AAA CTA GGA ACA TTG CCA
CCA GGG CTC ATA ACG TTC TAC AAT CTG ACA CAA CCA CTT CAG AGA AAA TGG CAC
TTA CTT GGA CTG GGT TAT GAT AAA GGA ATC GAT GTC AAG AAG ATT GAA AGA TCA
GCT GTT ATA CAT TAC AAT GGA CAC ATG AAA CCA TGG ACA GAG ATG GGG ATA AGC
AAG TAT CAG CCA TAT TGG ACG AAG TAC ACC AAT TTT GAC CAT CCT TAC ATC TTT
ACT TGC AGG CTG TTT GAG TGA

Amino Acid Sequence of Sequence #8:
Genebank ID# NP_182171
Positions 1-528.

MTDACCLKGNEDKMVPRFGHGTWIGKAFNDTPEMLHERSLRQEKRLERANELMNDDSLQKLETAAMARSRS
VDSAPLGNYTIWKNEYRRGKSFEDMLRLMQDQIIMARVYSGLAKFTNNLALHQEIETQLMKLAWEEESTDI
DQEQRVLDSIRDMGQILARAHEQLYECKLVTNKLRAMLQTVEDELENEQTYITFLTQLASKALPDAIHCLT
MRLNLEYHLLPLPMRNFPRRENLENPKLYHYALFSDNVLAASVVVNSTVMNAQDPSRHVFHLVTDKLNFGA
MSMWFLLNPPGEATIHVQRFEDFTWLNSSYSPVLSQLESAAMKKFYFKTARSESVESGSENLKYRYPKYMS
MLNHLRFYIPRIFPKLEKILFVDDDVVVQKDLTPLWSIDLKGKVNENFDPKFCGWAYGMNIFDLKEWKKNN
ITETYHFWQNLNENRTLWKLGTLPPGLITFYNLTQPLQRKWHLLGLGYDKGIDVKKIERSAVIHYNGHMKP
WTEMGISKYQPYWTKYTNFDHPYIFTCRLFE

[redacted]

Sequence #9

Gene name: At3g01040
GeneBank accession # for reference: NM_110969.1 GI:18395686

Nucleotide sequence of Sequence #9:
Positions 1-1533.

ATG AAG ATC AAA GTC GCA GCT CGT CAC ATC TCT TAC CGA ACT CTC TTC CAC ACT
ATC TTA ATC CTC GCT TTC TTG TTA CCT TTT GTT TTC ATC CTA ACC GCT GTT GTT
ACC CTT GAA GGT GTC AAC AAG TGC TCC TCT TTT GAT TGT TTC GGG AGG CGG CTA
GGA CCA CGT CTT CTT GGT AGG ATA GAT GAT TCA GAG CAG AGA CTA GTT AGA GAT
TTT TAC AAA ATT CTA AAT GAA GTA AGC ACT CAA GAA ATT CCA GAT GGT TTA AAG
CTT CCA GAG TCT TTT AGT CAA CTG GTT TCG GAT ATG AAG AAC AAC CAC TAT GAT
GCT AAA ACA TTT GCC CTC GTA TTT CGA GCT ATG GTA GAG AAG TTT GAA AGG GAT
TTA AGG GAA TCC AAA TTT GCA GAA CTC ATG AAC AAG CAC TTT GCT GCA AGT TCA
ATT CCA AAA GGA ATT CAC TGT CTC TCT TTA AGA CTA ACC GAT GAA TAT TCC TCC
AAT GCT CAT GCC CGG AGA CAG CTT CCT TCC CCG GAG CTT CTC CCT GTT CTC TCA
GAC AAT GCT TAC CAC CAT TTT GTT CTA GCT ACA GAT AAT ATC TTA GCT GCA TCG
GTT GTG GTC TCA TCT GCT GTT CAA TCA TCT TCA AAA CCC GAG AAA ATT GTC TTC
CAT GTT ATC ACA GAC AAG AAA ACC TAT GCG GGT ATG CAT TCT TGG TTT GCA CTC
AAT TCT GTT GCT CCT GCG ATT GTT GAA GTG AAA AGC GTT CAT CAG TTT GAT TGG
TTA ACA AGA GAG AAT GTT CCA GTT CTT GAA GCT GTG GAA AGC CAT AAC AGT ATC
AGA AAT TAT TAC CAT GGG AAT CAT ATT GCT GGT GCA AAC CTC AGC GAA ACA ACC
CCT CGA ACA TTT GCT TCG AAA CTG CAG TCA AGA AGT CCC AAA TAC ATA TCT TTG
CTC AAC CAT CTT AGA ATA TAT CTA CCA GAG CTT TTT CCG AAC TTA GAC AAG GTA
GTG TTC TTA GAT GAT GAT ATA GTG ATA CAG AAA GAT TTA TCT CCG CTT TGG GAT
ATT GAC CTT AAC GGG AAG GTT AAT GGA GCT GTG GAG ACT TGT CGA GGA GAA GAC
GTA TGG GTT ATG TCA AAG CGT CTT AGG AAC TAC TTC AAT TTT TCT CAC CCG CTC
ATC GCA AAG CAT TTA GAT CCC GAA GAA TGT GCT TGG GCT TAT GGA ATG AAT ATC
TTT GAT CTA CGG ACT TGG AGG AAG ACA AAT ATC AGA GAA ACG TAT CAT TCT TGG
CTT AAA GAG AAT CTG AAG TCG AAT CTA ACA ATG TGG AAA CTT GGA ACA TTG CCT
CCT GCT CTA ATA GCA TTT AAA GGT CAT GTT CAG CCA ATA GAT TCC TCT TGG CAT
ATG CTT GGA TTA GGT TAT CAG AGC AAG ACC AAC TTA GAA AAT GCG AAG AAA GCT
GCA GTG ATT CAT TAC AAT GGC CAA TCA AAG CCG TGG CTT GAG ATA GGT TTC GAG
CAT CTC AGA CCA TTC TGG ACA AAA TAT GTT AAC TAC TCC AAT GAT TTC ATT AAG
AAT TGT CAT ATC TTG GAA TAG

Amino Acid Sequence of Sequence #9:
Genebank ID# NP_186753
Positions 1-510.

MKIKVAARHISYRTLFHTILILAFLLPFVFILTAVVTLEGVNKCSSFDCFGRRLGPRLLGRIDDSEQRLVR
DFYKILNEVSTQEIPDGLKLPESFSQLVSDMKNNHYDAKTFALVFRAMVEKFERDLRESKFAELMNKHFAA
SSIPKGIHCLSLRLTDEYSSNAHARRQLPSPELLPVLSDNAYHHFVLATDNILAASVVVSSAVQSSSKPEK
IVFHVITDKKTYAGMHSWFALNSVAPAIVEVKSVHQFDWLTRENVPVLEAVESHNSIRNYYHGNHIAGANL
SETTPRTFASKLQSRSPKYISLLNHLRIYLPELFPNLDKVVFLDDDIVIQKDLSPLWDIDLNGKVNGAVET
CRGEDVWVMSKRLRNYFNFSHPLIAKHLDPEECAWAYGMNIFDLRTWRKTNIRETYHSWLKENLKSNLTMW
KLGTLPPALIAFKGHVQPIDSSWHMLGLGYQSKTNLENAKKAAVIHYNGQSKPWLEIGFEHLRPFWTKYVN
YSNDFIKNCHILE

Sequence #10

Gene name: At3g02350
GeneBank accession # for reference: AY056202
GI:15810326

Nucleotide sequence of Sequence #10:
Positions 243-1928.

ATG GCG GTG GCC TTC CGT GGA GGC CGG GGA GGC GTC GGA TCC GGC CAA TCT ACC
GGA CTT CGT AGT TTC TTC TCC TAC CGG ATC TTT ATC TCC GCT TTG TTC TCT TTT
CTC TTC CTC GCC ACT TTC TCC GTC GTT CTT AAC TCC TCT CGT CAT CAG CCT CAT
CAG GAT CAT ACA TTG CCG AGT ATG GGC AAC GCA TAT ATG CAG AGG ACG TTT TTG
GCT TTG CAA TCG GAT CCA TTG AAA ACT AGG TTG GAT CTG ATA CAC AAG CAA GCC
ATT GAT CAT TTG ACA CTG GTG AAT GCG TAT GCT GCT TAC GCT AGG AAG CTA AAG
CTT GAT GCT TCT AAG CAG CTT AAG CTC TTC GAA GAT TTG GCT ATC AAC TTC TCG
GAT TTG CAG TCG AAA CCT GGT TTG AAA TCT GCT GTG TCT GAT AAT GGT AAT GCT
CTT GAG GAG GAT TCG TTT AGG CAG CTT GAG AAA GAA GTG AAG GAT AAG GTG AAG
ACA GCG AGG ATG ATG ATC GTT GAG TCT AAA GAG AGT TAT GAT ACA CAG CTT AAA
ATC CAG AAG TTG AAA GAT ACA ATC TTT GCT GTC CAA GAA CAG TTG ACA AAG GCT
AAG AAA AAC GGT GCG GTT GCT AGC TTG ATT TCA GCC AAG TCG GTT CCT AAA AGT
CTT CAT TGT TTG GCC ATG AGG CTT GTA GGA GAG AGG ATC TCT AAT CCT GAG AAG
TAC AAG GAT GCT CCA CCT GAC CCA GCC GCA GAG GAT CCA ACT CTT TAC CAC TAT
GCG ATT TTC TCT GAT AAT GTC ATT GCT GTG TCT GTT GTG GTG AGA TCG GTT GTG
ATG AAC GCT GAG GAG CCA TGG AAG CAT GTC TTC CAT GTG GTG ACA GAT CGG ATG
AAT CTC GCA GCC ATG AAG GTG TGG TTT AAG ATG CGT CCT TTG GAC CGT GGT GCC
CAT GTT GAG ATT AAA TCC GTG GAG GAT TTC AAG TTC TTA AAC TCT TCC TAT GCG
CCG GTC TTG AGG CAG CTT GAG TCT GCC AAG TTG CAG AAG TTT TAC TTT GAG AAT
CAA GCT GAG AAC GCA ACT AAA GAT TCA CAT AAC CTC AAG TTC AAG AAC CCC AAG
TAT CTC TCG ATG TTG AAC CAT CTC AGA TTT TAC TTA CCA GAG ATG TAT CCG AAG
CTG AAT AAG ATT TTG TTC TTG GAC GAT GAT GTT GTG GTG CAG AAA GAC GTG ACT
GGT TTA TGG AAA ATC AAC TTG GAT GGC AAG GTG AAT GGA GCC GTT GAG ACA TGT
TTT GGT TCT TTT CAT CGA TAT GGT CAA TAC TTA AAC TTC TCT CAT CCT TTG ATC
AAA GAG AAC TTT AAC CCC AGT GCC TGT GCT TGG GCC TTT GGA ATG AAC ATA TTC
GAT CTC AAT GCC TGG AGA CGC GAG AAG TGC ACC GAT CAA TAC CAT TAC TGG CAG
AAC CTG AAT GAA GAC AGA ACT CTC TGG AAA TTG GGA ACT CTA CCT CCG GGA TTG
ATC ACA TTC TAT TCA AAG ACG AAA TCA TTG GAC AAA TCA TGG CAT GTA CTT GGG
TTA GGC TAT AAC CCG GGA GTG AGC ATG GAC GAA ATC AGA AAT GCA GGA GTG ATT
CAT TAC AAT GGA AAC ATG AAA CCG TGG CTA GAC ATT GCG ATG AAC CAA TAC AAG
TCT CTC TGG ACT AAA TAT GTT GAT AAC GAA ATG GAG TTT GTG CAG ATG TGC AAT
TTT GGT CTC TAA

[redacted]

Amino Acid Sequence of Sequence #10:
Genebank ID# NP_566170
Positions 1-561.

MAVAFRGGRGGVGSGQSTGLRSFFSYRIFISALFSFLFLATFSVVLNSSRHQPHQDHTLPSMGNAYMQRTF
LALQSDPLKTRLDLIHKQAIDHLTLVNAYAAYARKLKLDASKQLKLFEDLAINFSDLQSKPGLKSAVSDNG
NALEEDSFRQLEKEVKDKVKTARMMIVESKESYDTQLKIQKLKDTIFAVQEQLTKAKKNGAVASLISAKSV
PKSLHCLAMRLVGERISNPEKYKDAPPDPAAEDPTLYHYAIFSDNVIAVSVVVRSVVMNAEEPWKEVFHVV
TDRMNLAAMKVWFKMRPLDRGAHVEIKSVEDFKFLNSSYAPVLRQLESAKLQKFYFENQAENATKDSHNLK
FKNPKYLSMLNHLRFYLPEMYPKLNKILFLDDDVVVQKDVTGLWKINLDGKVNGAVETCFGSFHRYGQYLN
FSHPLIKENFNPSACAWAFGMNIFDLNAWRREKCTDQYHYWQNLNEDRTLWKLGTLPPGLITFYSKTKSLD
KSWHVLGLGYNPGVSMDEIRNAGVIHYNGNMKPWLDIAMNQYKSLWTKYVDNEMEFVQMCNFGL

Sequence #11

Gene name: at3g25140
GeneBank accession # for reference: NM_113418.2 GI:22331320

Nucleotide sequence of Sequence #11:
Positions 38-1717.

ATG GCT AAT CAC CAC CGA CTT TTA CGC GGC GGC GGA TCT CCG GCC ATA ATC GGT
GGC AGA ATC ACA CTC ACA GCT TTC GCT TCC ACT ATC GCA CTC TTC CTC TTC ACT
CTC TCC TTC TTC TTC GCT TCA GAT TCT AAC GAT TCT CCT GAT CTC CTT CTT CCC
GGT GTT GAG TAC TCT AAT GGA GTC GGA TCT AGA AGA TCC ATG TTG GAT ATC AAA
TCG GAT CCG CTT AAG CCA CGG TTG ATT CAG ATC CGG AAA CAA GCT GAT GAT CAT
CGG TCA TTA GCA TTA GCT TAT GCT TCT TAC GCG AGA AAG CTT AAG CTC GAG AAT
TCG AAA CTC GTC AGG ATC TTC GCT GAT CTT TCG AGG AAT TAC ACG GAT CTG ATT
AAC AAA CCG ACG TAT CGA GCT TTG TAT GAT TCT GAT GGA GCC TCG ATT GAA GAA
TCT GTG CTT AGG CAA TTT GAG AAA GAA GTT AAG GAA CGG ATT AAA ATG ACT CGT
CAA GTG ATT GCT GAA GCT AAA GAG TCT TTT GAT AAT CAG TTG AAG ATT CAG AAG
CTG AAA GAT ACG ATT TTC GCT GTT AAC GAA CAG TTA ACT AAT GCT AAG AAG CAA
GGT GCG TTT TCG AGT TTG ATC GCT GCG AAA TCG ATT CCG AAA GGA TTG CAT TGT
CTT GCT ATG AGG CTG ATG GAA GAG AGG ATT GCT CAC CCT GAG AAG TAT ACT GAT
GAA GGG AAA GAT AGA CCG CGG GAG CTC GAG GAT CCG AAT CTT TAC CAT TAC GCT
ATA TTT TCG GAT AAT GTG ATT GCG GCT TCG GTG GTT GTG AAC TCT GCT GTG AAG
AAT GCT AAG GAG CCG TGG AAG CAT GTT TTT CAC GTT GTG ACT GAT AAG ATG AAT
CTT GGA GCT ATG CAG GTT ATG TTT AAA CTG AAG GAG TAT AAA GGA GCT CAT GTA
GAA GTT AAA GCT GTT GAG GAT TAT ACG TTT TTG AAC TCT TCG TAT GTG CCT GTG
TTG AAG CAG TTA GAA TCT GCG AAT CTT CAG AAG TTT TAT TTC GAG AAT AAG CTC
GAG AAT GCG ACG AAA GAT ACC ACG AAT ATG AAG TTC AGG AAC CCC AAG TAT TTA
TCT ATA TTG AAT CAC TTG AGG TTT TAT TTA CCC GAG ATG TAC CCG AAA CTA CAT
AGG ATA CTG TTT TTG GAC GAT GAT GTG GTT GTG CAG AAG GAT TTA ACG GGT CTG
TGG GAG ATT GAT ATG GAT GGG AAA GTG AAT GGA GCT GTA GAG ACT TGT TTT GGG
TCG TTT CAT CGG TAC GCT CAA TAC ATG AAT TTC TCA CAT CCT TTG ATC AAA GAG
AAG TTT AAT CCC AAA GCA TGT GCG TGG GCG TAT GGA ATG AAC TTC TTT GAT CTT
GAT GCT TGG AGA AGA GAG AAG TGC ACA GAA GAA TAT CAC TAC TGG CAA AAT CTG
AAC GAG AAC AGG GCT CTA TGG AAA CTG GGG ACG TTA CCA CCG GGA CTG ATC ACC
TTT TAC TCA ACC ACA AAG CCG CTG GAC AAA TCA TGG CAT GTG CTT GGG CTG GGT
TAC AAT CCG AGC ATT AGC ATG GAT GAG ATC CGC AAC GCT GCA GTG GTA CAC TTC
AAC GGT AAC ATG AAG CCA TGG CTT GAC ATA GCT ATG AAC CAG TTT CGA CCA CTT
TGG ACC AAA CAC GTC GAC TAT GAC CTC GAG TTT GTT CAG GCT TGC AAT TTT GGC
CTC TGA

Amino Acid Sequence of Sequence #11:
Genebank ID# NP_189150
Positions 1-559.

MANHHRLLRGGGSPAIIGGRITLTAFASTIALFLFTLSFFFASDSNDSPDLLLPGVEYSNGVGSRRSMLDI
KSDPLKPRLIQIRKQADDHRSLALAYASYARKLKLENSKLVRIFADLSRNYTDLINKPTYRALYDSDGASI
EESVLRQFEKEVKERIKMTRQVIAEAKESFDNQLKIQKLKDTIFAVNEQLTNAKKQGAFSSLIAAKSIPKG
LHCLAMRLMEERIAHPEKYTDEGKDRPRELEDPNLYHYAIFSDNVIAASVVVNSAVKNAKEPWKEVFHVVT
DKMNLGAMQVMFKLKEYKGAHVEVKAVEDYTFLNSSYVPVLKQLESANLQKFYFENKLENATKDTTNMKFR
NPKYLSILNHLRFYLPEMYPKLHRILFLDDDVVVQKDLTGLWEIDMDGKVNGAVETCFGSFHRYAQYMNFS
HPLIKEKFNPKACAWAYGMNFFDLDAWRREKCTEEYHYWQNLNENRALWKLGTLPPGLITFYSTTKPLDKS
WHVLGLGYNPSISMDEIRNAAVVHFNGNMKPWLDIAMNQFRPLWTKHVDYDLEFVQACNFGL

Sequence #12

Gene name: At3g58790
GeneBank accession # for reference: NM_115741.2 GI:22331856

Nucleotide sequence of Sequence #12:
Positions 353-1975.

ATG AAG TTT TAC ATA TCA GCG ACG GGG ATT AAG AAG GTT ACG ATA TCA AAT CCC
GGC GTC GGA ATC GGT AAA GGA AGC GGA GGA TGT GCG GCT GCA GCG GCG GCG TTA
GCA GCG CGG AGA TTC TCT AGT CGC ACG TTG TTA CTG TTG CTG CTG CTG CTC GCT
ATC GTC CTC CCT TTT ATC TTC GTC AGG TTC GCG TTT CTC GTC CTC GAA TCT GCC
TCC GTT TGC GAT TCA CCA CTC GAT TGC ATG GGA CTC AGA CTT TTC CGT GGG GGC
GAC ACA TCT CTG AAA ATT GGG GAA GAG TTG ACA CGG GCT CTA GTG GAA GAG ACG
ACA GAT CAT CAG GAC GTT AAT GGA AGA GGA ACG AAG GGA TCA TTG GAG TCA TTC
GAC GAC CTT GTT AAG GAG ATG ACG TTA AAA CGC CGT GAC ATA AGG GCG TTT GCT
TCC GTG ACT AAG AAG ATG CTG TTG CAG ATG GAA CGT AAA GTC CAA TCA GCG AAA
CAT CAT GAG TTA GTG TAC TGG CAT TTA GCC TCT CAC GGT ATT CCT AAA AGC CTC
CAT TGC CTT TCC CTC AGA TTA ACT GAA GAG TAC TCT GTA AAT GCA ATG GCT CGA
ATG CGT TTG CCT CCG CCT GAG TCC GTA TCA CGT CTG ACC GAC CCA TCT TTT CAT
CAT ATT GTC CTC CTG ACT GAC AAT GTC CTT GCT GCC TCT GTC GTC ATA TCG TCT
ACT GTA CAA AAC GCT GTG AAT CCC GAG AAG TTT GTC TTT CAT ATT GTT ACC GAT
AAG AAA ACC TAT ACC CCT ATG CAT GCT TGG TTT GCT ATC AAC TCT GCT TCA TCA
CCA GTT GTT GAA GTA AAG GGA CTT CAT CAG TAT GAT TGG CCT CAA GAA GTG AAC
TTC AAA GTT AGA GAG ATG CTG GAC ATT CAC CGC TTA ATT TGG AGA CGA CAT TAT
CAA AAT TTG AAA GAC TCT GAT TTT AGT TTT GTT GAG GGT ACT CAT GAG CAG TCC
TTG CAA GCT CTA AAT CCT AGC TGC CTT GCC CTT TTG AAC CAT CTT CGC ATT TAC
ATT CCC AAG CTT TTT CCA GAT CTC AAC AAG ATA GTG TTG TTG GAT GAT GAT GTA
GTA GTA CAG AGC GAT CTT TCG TCT TTA TGG GAA ACG GAT CTC AAC GGT AAA GTT
GTT GGT GCT GTC GTT GAT TCG TGG TGC GGA GAC AAC TGT TGC CCC GGA AGA AAA
TAC AAA GAC TAT TTC AAC TTC TCA CAT CCT TTG ATC TCA TCA AAC TTA GTT CAA
GAA GAC TGT GCT TGG CTT TCT GGT ATG AAT GTC TTT GAT CTC AAA GCC TGG AGA
CAA ACC AAT ATT ACT GAA GCT TAC TCT ACA TGG CTA AGA CTC AGT GTT AGG TCA
GGA CTA CAA TTA TGG CAA CCA GGG GCT TTA CCA CCG ACA TTA CTT GCT TTC AAA
GGA CTT ACA CAG TCT CTT GAA CCA TCA TGG CAC GTC GCT GGA CTA GGT TCT CGA
TCC GTA AAA TCC CCT CAA GAG ATT CTG AAA TCT GCT TCG GTT TTA CAT TTC AGC
GGT CCA GCA AAA CCG TGG CTA GAG ATC AGT AAC CCT GAG GTA CGA TCT CTT TGG
TAT AGA TAC GTA AAT TCC TCC GAC ATC TTC GTT AGA AAA TGC AAA ATC ATG AAC
TGA

Amino Acid Sequence of Sequence #12:
Genebank ID# NP_191438
Positions 1-540.

MKFYISATGIKKVTISNPGVGIGKGSGGCAAAAAALAARRFSSRTLLLLLLLLAIVLPFIFVRFAFLVLES
ASVCDSPLDCMGLRLFRGGDTSLKIGEELTRALVEETTDHQDVNGRGTKGSLESFDDLVKEMTLKRRDIRA
FASVTKKMLLQMERKVQSAKHHELVYWHLASHGIPKSLHCLSLRLTEEYSVNAMARMRLPPPESVSRLTDP
SFHHIVLLTDNVLAASVVISSTVQNAVNPEKFVFHIVTDKKTYTPMHAWFAINSASSPVVEVKGLHQYDWP
QEVNFKVREMLDIHRLIWRRHYQNLKDSDFSFVEGTHEQSLQALNPSCLALLNHLRIYIPKLFPDLNKEVL
LDDDVVVQSDLSSLWETDLNGKVVGAVVDSWCGDNCCPGRKYKDYFNFSHPLISSNLVQEDCAWLSGMSVF
DLKAWRQTNITEAYSTWLRLSVRSGLQLWQPGALPPTLLAFKGLTQSLEPSWHVAGLGSRSVKSPQEILKS
ASVLHFSGPAKPWLEISNPEVRSLWYRYVNSSDIFVRKCKIMN

Sequence #13

Gene name: At4g48270
GeneBank accession # for reference: NM_119989.1 GI:18420208

Nucleotide sequence of Sequence #13:
Positions 1-1977.

ATG TCG ACG ATT TGC AGC CAC AGA GAA CTT AAA GCT TAT CGT CCG CTG CAA GAT
AAT AAT CTA CAG GAG GTG TAT GCT TCC TCA GCT GCT GCA GTG CAC TAC GAT CCA
GAT CTG AAA CTG TTA TCT CAG GAT GTG AAC ATA GTT GCG ACA TAC AGT GAC CAT
TAC GGC AAT ATA CGC CTT GGT AGG GTG AAA ATG GGG GAT CTT TCA CCT TCT TGG
GTT TTG GAG AAT CCT GCC TAT CAA GTT AGC CGC AAA ACA AAA GGT TCG CAG CTA
GTT ATA CCA CGG GAT TCA TTT CAA AAT GAT ACT GGA ATG GAA GAT AAT GCA AGC
CAT TCT ACA ACT AAT CAG ACT GAT GAA AGC GAA AAT CAG TTT CCA AAC GTG GAT
TTT GCA AGC CCA GCA AAA CTG AAG CGG CAG ATT TTA CGT CAG GAA AGG AGA GGT
CAA CGA ACT TTA GAG CTG ATC CGA CAA GAA AAG GAA ACT GAT GAG CAG ATG CAA
GAA GCA GCC ATT CAG AAG TCA ATG AGC TTT GAA AAC TCA GTC ATA GGG AAA TAC
AGT ATA TGG AGG AGA GAC TAT GAG AGC CCA AAT GCT GAT GCT ATC TTG AAG CTT
ATG AGA GAC CAG ATC ATA ATG GCA AAA GCA TAT GCA AAT ATT GCC AAA TCA AAA
AAT GTA ACC AAT CTG TAC GTT TTC TTG ATG CAG CAG TGT GGA GAA AAT AAA CGT
GTT ATA GGT AAA GCA ACC TCT GAT GCT GAC CTT CCT TCA AGC GCT CTT GAT CAA
GCA AAA GCC ATG GGC CAT GCA CTC TCT CTT GCA AAA GAC GAG TTA TAT GAC TGC
CAT GAA CTT GCA AAA AAG TTC CGG GCC ATC CTT CAG TCC ACT GAA CGC AAA GTA
GAT GGA CTG AAG AAA AAG GGA ACC TTC TTA ATT CAG CTA GCT GCC AAA ACA TTT
CCC AAG CCA TTG CAT TGC CTG AGT CTG CAG CTA GCG GCA GAC TAT TTT ATT CTA
GGT TTC AAT GAA GAG GAT GCA GTG AAA GAG GAT GTC AGT CAA AAG AAG CTT GAA
GAT CCT TCG CTC TAT CAC TAT GCG ATC TTT TCG GAT AAC GTT CTG GCT ACA TCA
GTG GTG GTG AAC TCC ACT GTC TTG AAT GCA AAG GAA CCG CAG AGG CAT GTG TTC
CAT ATA GTA ACT GAC AAA CTG AAT TTT GGT GCA ATG AAG ATG TGG TTT CGC ATC
AAT GCT CCT GCT GAT GCG ACG ATT CAA GTT GAA AAC ATA AAT GAT TTC AAG TGG
CTG AAC TCC TCT TAC TGC TCT GTT CTA CGG CAG CTT GAA TCT GCA AGG CTG AAA
GAA TAC TAT TTC AAA GCA AAT CAT CCT TCA TCA ATC TCA GCT GGC GCA GAT AAT
CTA AAG TAC CGC AAC CCA AAG TAT CTA TCG ATG CTG AAT CAT CTC AGA TTC TAC
CTT CCT GAG GTT TAT CCG AAG CTG GAG AAG ATT CTG TTT CTA GAC GAT GAC ATT
GTG GTG CAG AAG GAC CTG GCA CCA CTA TGG GAA ATA GAC ATG CAA GGA AAA GTG
AAT GGT GCG GTG GAG ACG TGC AAG GAG AGC TTC CAC AGA TTT GAC AAG TAC CTC
AAC TTC TCA AAT CCA AAG ATT TCA GAG AAT TTT GAC GCT GGT GCT TGT GGG TGG
GCA TTT GGG ATG AAT ATG TTT GAC CTG AAA GAG TGG AGG AAA CGG AAC ATT ACA
GGWUUUUUUUUUUUÕ•w[ÿÒ¯ÿÿÿCAA GAC TTG AAT GAA GAC AGA ACA CTG TGG AAG CTG
GGA TCG TTG CCA CCG GGG CTG ATA ACA TTT TAC AAC CTG ACG TAT GCA ATG GAT
AGG AGC TGG CAC GTA CTA GGG CTG GGA TAT GAC CCA GCG CTA AAC CAA ACA GCA
ATA GAG AAT GCA GCG GTA GTG CAT TAC AAT GGG AAC TAC AAG CCA TGG CTG GGT
TTA GCA TTC GCC AAG TAC AAA CCG TAC TGG TCC AAG TAC GTT GAG TAC GAC AAC
CCT TAT CTC CGA CGG TGC GAC ATC AAT GAA TGA

Amino Acid Sequence of Sequence #13:
Genebank ID# NP_195540
Positions 1-658.

MSTICSHRELKAYRPLQDNNLQEVYASSAAAVHYDPDLKLLSQDVNIVATYSDHYGNIRLGRVKMGDLSPS
WVLENPAYQVSRKTKGSQLVIPRDSFQNDTGMEDNASHSTTNQTDESENQFPNVDFASPAKLKRQILRQER
RGQRTLELIRQEKETDEQMQEAAIQKSMSFENSVIGKYSIWRRDYESPNADAILKLMRDQIIMAKAYANIA
KSKNVTNLYVFLMQQCGENKRVIGKATSDADLPSSALDQAKAMGHALSLAKDELYDCHELAKKFRAILQST
ERKVDGLKKKGTFLIQLAAKTFPKPLHCLSLQLAADYFILGFNEEDAVKEDVSQKKLEDPSLYHYAIFSDN
VLATSVVVNSTVLNAKEPQRHVFHIVTDKLNFGAMKMWFRINAPADATIQVENINDFKWLNSSYCSVLEQL
ESARLKEYYFKANHPSSISAGADNLKYRNPKYLSMLNHLRFYLPEVYPKLEKILFLDDDIVVQKDLAPLWE
IDMQGKVNGAVETCKESFHRFDKYLNFSNPKISENFDAGACGWAFGMNMFDLKEWRKRNITGIYHYWQDLN
EDRTLWKLGSLPPGLITFYNLTYAMDRSWHVLGLGYDPALNQTAIENAAVVHYNGNYKPWLGLAFAKYKPY
WSKYVEYDNPYLRRCDINE

Sequence #14

Gene name: At5g15470
GeneBank accession # for reference: AY056202
GI:15810326

Nucleotide sequence of Sequence #14:
Positions 1-1719.

ATG CGG TCT GGG AGA AGA CCA CAA GGA AGA AGA ATA GCT ATA AGG AAC GAA ACA
GAG ACG GAA TTG AGA TCG AGA ATC GGA GAG AAA TCG GGA TCG GAG AAG CGG AAT
CTG ATC GGA TCT GAT GTT GGG ATG CAG CTT CAC ATA TCG CCG AGT ATG AGA AGC
ATT ACG ATT TCG AGC AGC AAT GAG TTT ATT GAC TTG ATG AAG ATC AAG GTC GCA
GCT CGT CAC ATC TCT TAC CGA ACT CTC TTC CAC ACC ATC TTA ATC CTC GCT TTC
TTG TTG CCT TTT GTT TTC ATT CTC ACC GCT GTT GTT ACC CTT GAG GGT GTC AAC
AAA TGC TCC TCC ATT GGG AGG CGG ATA GGT CCA CGT CTT CTT GGT AGG GTA GAT
GAT TCA GAG AGA CTA GCT AGA GAC TTT TAT AAA ATT CTA AAC GAA GTA AGC ACT
CAA GAA ATT CCA GAT GGT TTG AAG CTT CCA AAT TCT TTT AGT CAA CTT GTT TCC
GAT ATG AAG AAT AAC CAC TAT GAT GCA AAA ACA TTT GCT CTT GTG CTG CGA GCC
ATG ATG GAG AAG TTT GAA CGT GAT ATG AGG GAA TCG AAA TTT GCA GAA CTT ATG
AAC AAG CAC TTT GCA GCA AGT TCC ATT CCC AAA GGC ATT CAT TGT CTC TCT CTA
AGA CTG ACA GAT GAA TAT TCC TCC AAT GCT CAT GCT CGA AGA CAG CTT CCT TCA
CCA GAG TTT CTC CCT GTT CTT TCA GAT AAT GCT TAC CAC CAC TTT ATT TTG TCC
ACG GAC AAT ATT TTG GCT GCC TCA GTT GTG GTC TCA TCC GCT GTT CAG TCA TCT
TCA AAA CCC GAG AAA ATT GTC TTT CAC ATC ATT ACA GAC AAG AAA ACC TAT GCG
GGT ATG CAT TCA TGG TTT GCG CTT AAT TCT GTT GCA CCA GCA ATT GTT GAG GTT
AAA GGT GTT CAT CAG TTT GAC TGG TTG ACG AGA GAG AAT GTT CCG GTT TTG GAA
GCT GTG GAA AGC CAT AAT GGT GTC AGG GAC TAT TAT CAT GGG AAT CAT GTC GCT
GGG GCA AAC CTC ACC GAA ACA ACT CCT CGA ACA TTT GCT TCA AAA TTG CAG TCT
AGA AGT CCA AAA TAC ATA TCT TTG CTC AAC CAT CTT AGA ATA TAT ATA CCA GAG
CTT TTC CCG AAC TTG GAC AAG GTG GTT TTC TTA GAC GAT GAT ATA GTT GTC CAG
GGA GAC TTA ACT CCA CTT TGG GAT GTT GAC CTC GGT GGT AAG GTC AAT GGG GCA
GTA GAG ACT TGC AGG GGT GAA GAT GAA TGG GTG ATG TCA AAG CGT TTA AGG AAC
TAC TTC AAT TTC TCT CAC CCG CTC ATC GCA AAG CAT TTA GAT CCT GAA GAA TGT
GCT TGG GCA TAT GGT ATG AAT ATC TTC GAT CTA CAA GCT TGG AGG AAA ACA AAT
ATC AGA GAA ACG TAT CAC TCT TGG CTT AGA GAG AAT CTA AAG TCA AAT CTG ACA
ATG TGG AAA CTT GGA ACC TTG CCT CCT GCT CTT ATC GCG TTC AAG GGT CAC GTA
CAC ATA ATA GAC TCG TCA TGG CAT ATG CTA GGA TTA GGC TAC CAG AGC AAG ACC
AAC ATA GAA AAT GTG AAG AAA GCA GCA GTG ATC CAC TAC AAT GGG CAG TCA AAG
CCA TGG CTG GAG ATT GGT TTC GAG CAT CTG CGG CCA TTC TGG ACC AAA TAC GTC
AAC TAC TCA AAT GAT TTC ATC AAG AAC TGT CAC ATA TTG GAG TAG

Amino Acid Sequence of Sequence #14:
Genebank ID# NP_197051
Positions 1-572.

MRSGRRPQGRRIAIRNETETELRSRIGEKSGSEKRNLIGSDVGMQLHISPSMRSITISSSNEFIDLMKKV
AARHISYRTLFHTILILAFLLPFVFILTAVVTLEGVNKCSSIGRRIGPRLLGRVDDSERLARDFYKILNEV
STQEIPDGLKLPNSFSQLVSDMKNNHYDAKTFALVLRAMMEKFERDMRESKFAELMNKHFAASSIPKGIHC
LSLRLTDEYSSNAHARRQLPSPEFLPVLSDNAYHHFILSTDNILAASVVVSSAVQSSSKPEKIVFHIIDDK
KTYAGMHSWFALNSVAPAIVEVKGVHQFDWLTRENVPVLEAVESHNGVRDYYHGNHVAGANLTETTPRIFA
SKLQSRSPKYISLLNHLRIYIPELFPNLDKVVFLDDDIVVQGDLTPLWDVDLGGKVNGAVETCRGEDRGVM
SKRLRNYFNFSHPLIAKHLDPEECAWAYGMNIFDLQAWRKTNIRETYHSWLRENLKSNLTMWKLGTLPPAL
IAFKGHVHIIDSSWHMLGLGYQSKTNIENVKKAAVIHYNGQSKPWLEIGFEHLRPFWTKYVNYSNDFINC
HILE

Sequence #15

Gene name: At5g54690
GeneBank accession # for reference: AY056202
GI:15810326.

Nucleotide sequence of Sequence #15:
Positions 1-1608.

ATG CAG TTA CAT ATA TCT CCG AGC TTG AGA CAT GTG ACT GTG GTC ACA GGG AAA
GGA TTG AGA GAG TTC ATA AAA GTT AAG GTT GGT TCT AGA AGA TTC TCT TAT CAA
ATG GTG TTT TAC TCT CTA CTC TTC TTC ACT TTT CTT CTC CGA TTC GTC TTT GAT
CTC TCC ACC GTT GAT ACT ATC GAC GGC GAT CCC TCT CCT TGC TCC TCT CTT GCT
TGC TTG GGG AAA AGA CTA AAG CCA AAG CTT TTA GGA AGA AGG GTT GAT TCT GAT
AAT GTT CCA GAA GCT ATG TAC CAA GTT TTA GAA CAG CCT TTA AGC GAA CAA GAA
CTC AAA GGA AGA TCA GAT ATA CCT CAA ACA CTT CAA GAT TTC ATG TCT GAA GTC
AAA AGA AGC AAA TCA GAC GCA AGA GAA TTT GCT CAA AAG CTA AAA GAA ATG GTG
ACA TTG ATG GAA CAG AGA ACA AGA ACG GCT AAG ATT CAA GAG TAT TTA TAT CGA
CAT GTC GCA TCA AGC AGC ATA CCG AAA CAA CTT CAC TGT TTA GCT CTT AAA CTA
GCC AAC GAA CAC TCG ATA AAC GCA GCG GCG CGT CTC CAG CTT CCA GAA GCT GAG
CTT GTC CCT ATG TTG GTA GAC AAC AAC TAC TTT CAC TTT GTC TTG GCT TCA GAC
AAT ATT CTT GCA GCT TCG GTT GTG GCT AAG TCG TTG GTT CAA AAT GCT TTA AGA
CCT CAT AAG ATC GTT CTT CAC ATC ATA ACG GAT AGG AAA ACT TAT TTC CCA ATG
CAA GCT TGG TTC TCA TTG CAT CCT CTG TCT CCA GCA ATA ATT GAG GTC AAG GCT
TTG CAT CAT TTC GAT TGG TTA TCG AAA GGT AAA GTA CCC GTT TTG GAA GCT ATG
GAG AAA GAT CAG AGA GTG AGG TCT CAA TTC AGA GGT GGA TCA TCG GTT ATT GTG
GCT AAT AAC AAA GAG AAC CCG GTT GTT GTT GCT GCT AAG TTA CAA GCT CTC AGC
CCT AAA TAC AAC TCC TTG ATG AAT CAC ATC CGT ATT CAT CTA CCA GAG TTG TTT
CCA AGC TTA AAC AAG GTT GTG TTT CTA GAC GAT GAC ATT GTG ATC CAA ACT GAT
CTT TCA CCT CTT TGG GAC ATT GAC ATG AAT GGA AAA GTA AAT GGA GCA GTG GAA
ACA TGT AGA GGA GAA GAC AAG TTT GTG ATG TCA AAG AAG TTC AAG AGT TAC CTC
AAC TTC TCG AAT CCG ACA ATT GCC AAA AAC TTC AAT CCA GAG GAA TGT GCA TGG
GCT TAT GGA ATG AAT GTT TTC GAC CTA GCG GCT TGG AGG AGG ACT AAC ATA AGC
TCC ACT TAC TAT CAT TGG CTT GAC GAG AAC TTA AAA TCA GAC CTG AGT TTG TGG
CAG CTG GGA ACT TTG CCT CCT GGG CTG ATT GCT TTC CAC GGT CAT GTC CAA ACC
ATA GAT CCG TTC TGG CAT ATG CTT GGT CTC GGA TAC CAA GAG ACC ACG AGC TAT
GCC GAT GCT GAA AGT GCC GCT GTT GTT CAT TTC AAT GGA AGA GCT AAG CCT TGG
CTG GAT ATA GCA TTT CCT CAT CTA CGT CCT CTC TGG GCT AAG TAT CTT GAT TCT
TCT GAC AGA TTT ATC AAG AGC TGT CAC ATT AGA GCA TCA TGA

Amino Acid Sequence of Sequence #15:
Genebank ID# NP_200280
Positions 1-535.

MQLHISPSLRHVTVVTGKGLREFIKVKVGSRRFSYQMVFYSLLFFTFLLRFVFVLSTVDTIDGDPSPCSSE
ACLGKRLKPKLLGRRVDSGNVPEAMYQVLEQPLSEQELKGRSDIPQTLQDFMSEVKRSKSDAREFAQKLKE
MVTLMEQRTRTAKIQEYLYRHVASSSIPKQLHCLALKLANEHSINAAARLQLPEAELVPMLVDNNYFHFWE
ASDNILAASVVAKSLVQNALRPHKIVLHIITDRKTYFPMQAWFSLHPLSPAIIEVKALHHFDWLSKGKVPW
LEAMEKDQRVRSQFRGGSSVIVANNKENPVVVAAKLQALSPKYNSLMNHIRIHLPELFPSLNKVVFLDDDE
VIQTDLSPLWDIDMNGKVNGAVETCRGEDKFVMSKKFKSYLNFSNPTIAKNFNPEECAWAYGMNVFDLAAW
RRTNISSTYYHWLDENLKSDLSLWQLGTLPPGLIAFHGHVQTIDPFWHMLGLGYQETTSYADAESAAVVHF
NGRAKPWLDIAFPHLRPLWAKYLDSSDRFIKSCHIRAS

APPENDIX II

Indx	Template	Bx	Primer	Box	MGIF INDEX	Cost	GEL
DNASEQL6	J533.6P Plasmid DS:GC Rich	58L 36	PEFO	58L 39	324192	25.00	N557/1
	J536.7P Plasmid DS:GC Rich	58L 37	PEFO	58L 39	324193	25.00	N557/2
	J536L.3P Plasmid DS:GC Rich	58L 38	PEFO	58L 39	324194	25.00	N557/3
	J533.6P Plasmid DS:GC Rich	58L 36	PERE	58L 40	324195	25.00	N557/4
	J536.7P Plasmid DS:GC Rich	58L 37	PERE	58L 40	324196	25.00	N557/5
	J536L.3P Plasmid DS:GC Rich	58L 38	PERE	58L 40	324197	25.00	N557/6
	J533.6P Plasmid DS:GC Rich	58L 36	331P	58L 41	324198	25.00	N557/7
	J536.7P Plasmid DS:GC Rich	58L 37	361P	58L 42	324199	25.00	N557/8
	J536L.3P Plasmid DS:GC	58L 38	6LIP	58L 43	324200	25.00	N557/9

DNASEQL6	Rich							

JASON STERLINE **Grand Total: 225.00**

Sequencing and Synthesis FacilityWork Order: 202019952

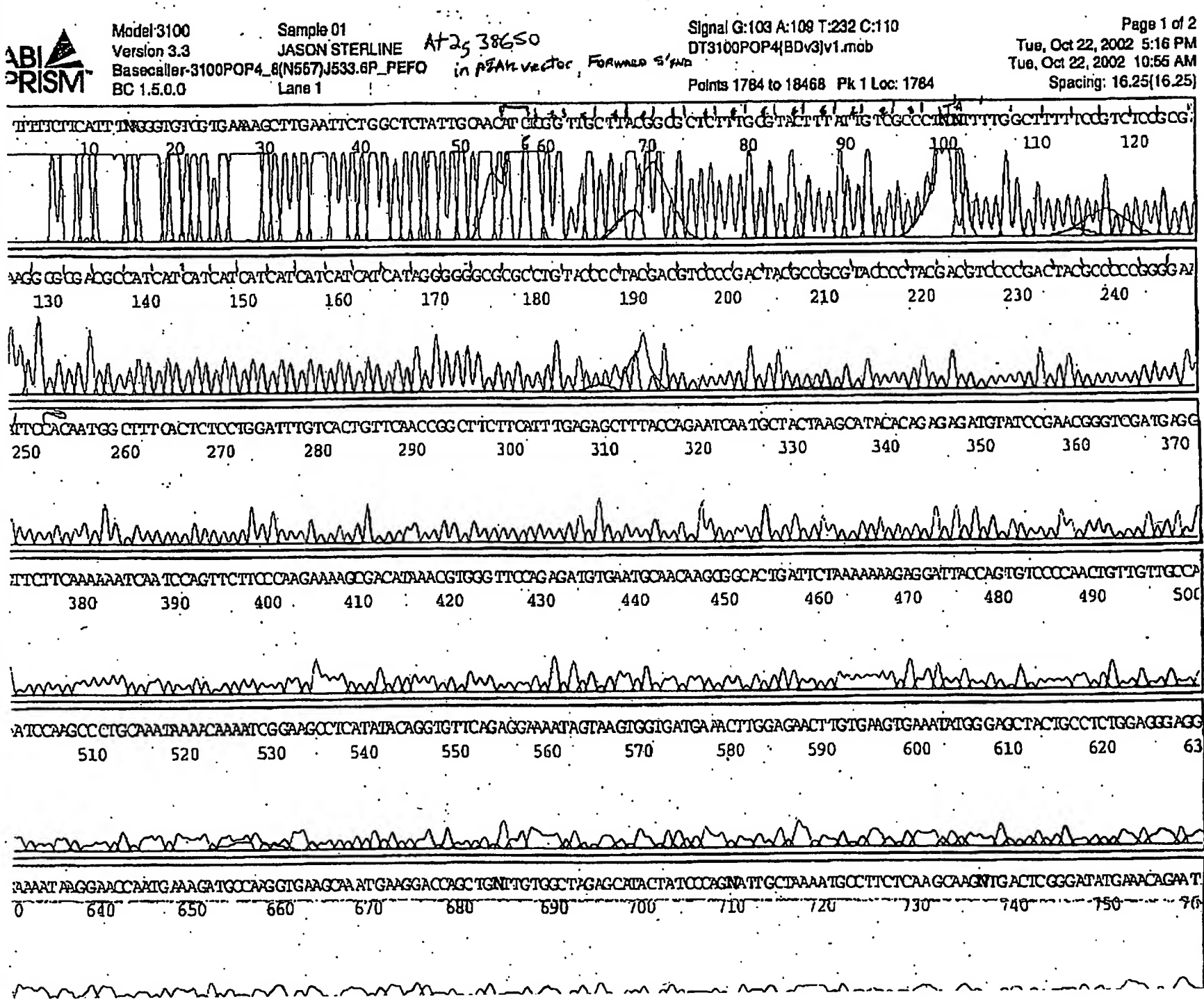
Comments:

Processed By: _____ Date: _____

Received By: _____ Date: _____

N557

*Failed samples may be in process for rerun at no additional charg
call MGIF for details. Chromatograms of failed samples are enclose
for your review (policy started 9/22/2000). You may also
retrieve past failed chromatograms from our web site.
Some past chromatograms may not be available on line.


ABI PRISM
Model 3100
Version 3.3
Basecaller-3100POP4_6(N567)J533.6P_PEFO
BC 1.5.0.0
Sample 01
JASON STERLINE
Lane 1
At2g 38650
in pZAN vector, Formed 5'end
Signal G:103 A:109 T:232 C:110
DT3100POP4(BDv3)v1.mob
Points 1784 to 18468 Pk 1 Loc: 1784
Page 1 of 2
Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.25(16.25)

ABI PRISM

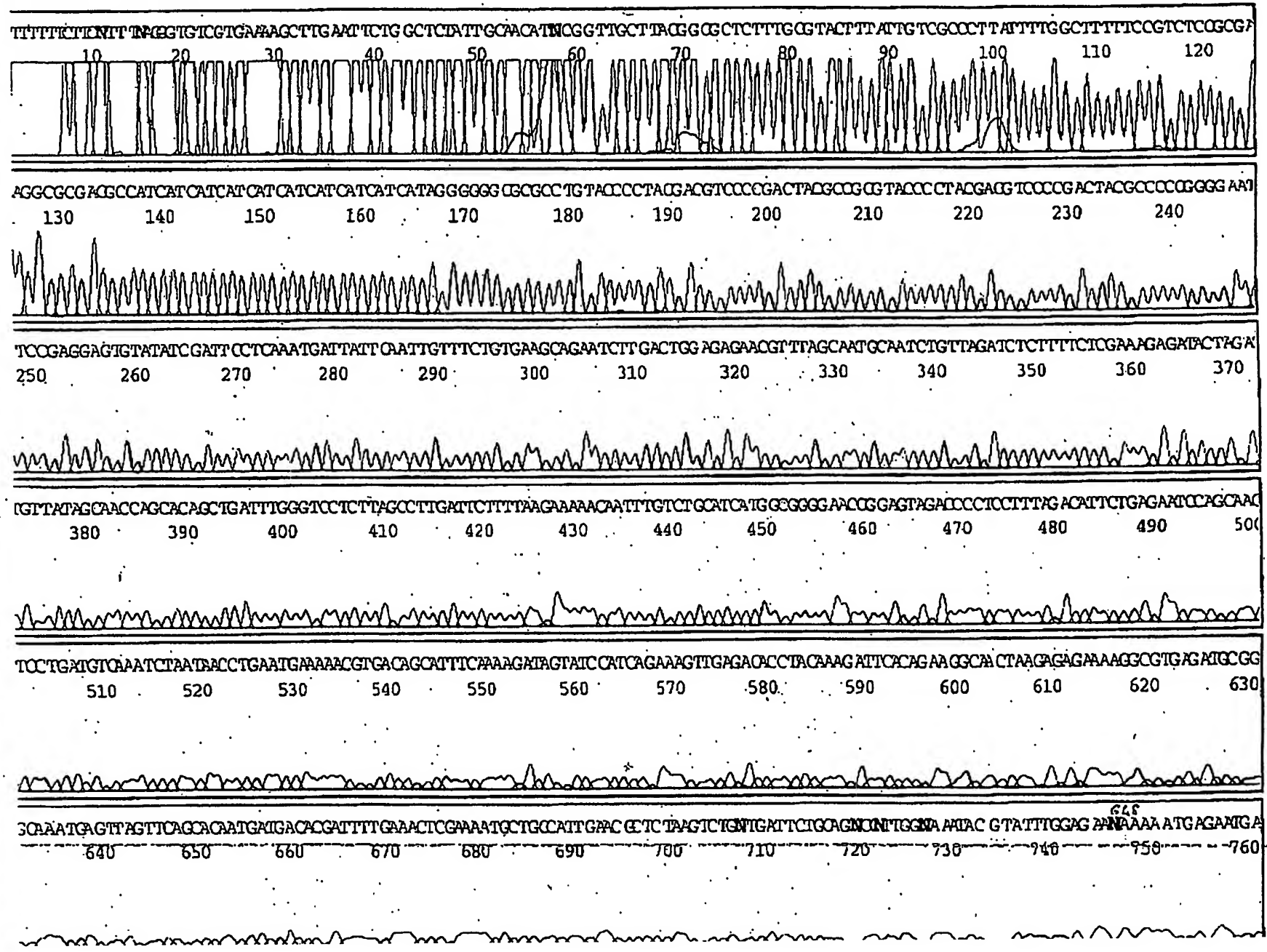
Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J536.7P_PEFQ
BC 1.5.0.0

Sample 02
JASON STERLINE
Lane 3

At 3₅ 61130
in PEAK, FORWARD FROM 5'END

Signal G:135 A:158 T:305 C:124
DT3100POP4(BDv3)v1.mob
Points 1774 to 18468 Pk 1 Loc: 1774

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.25(16.25)



ABI PRISM

Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J536L.3P_PEFO
BC 1.5.0.0

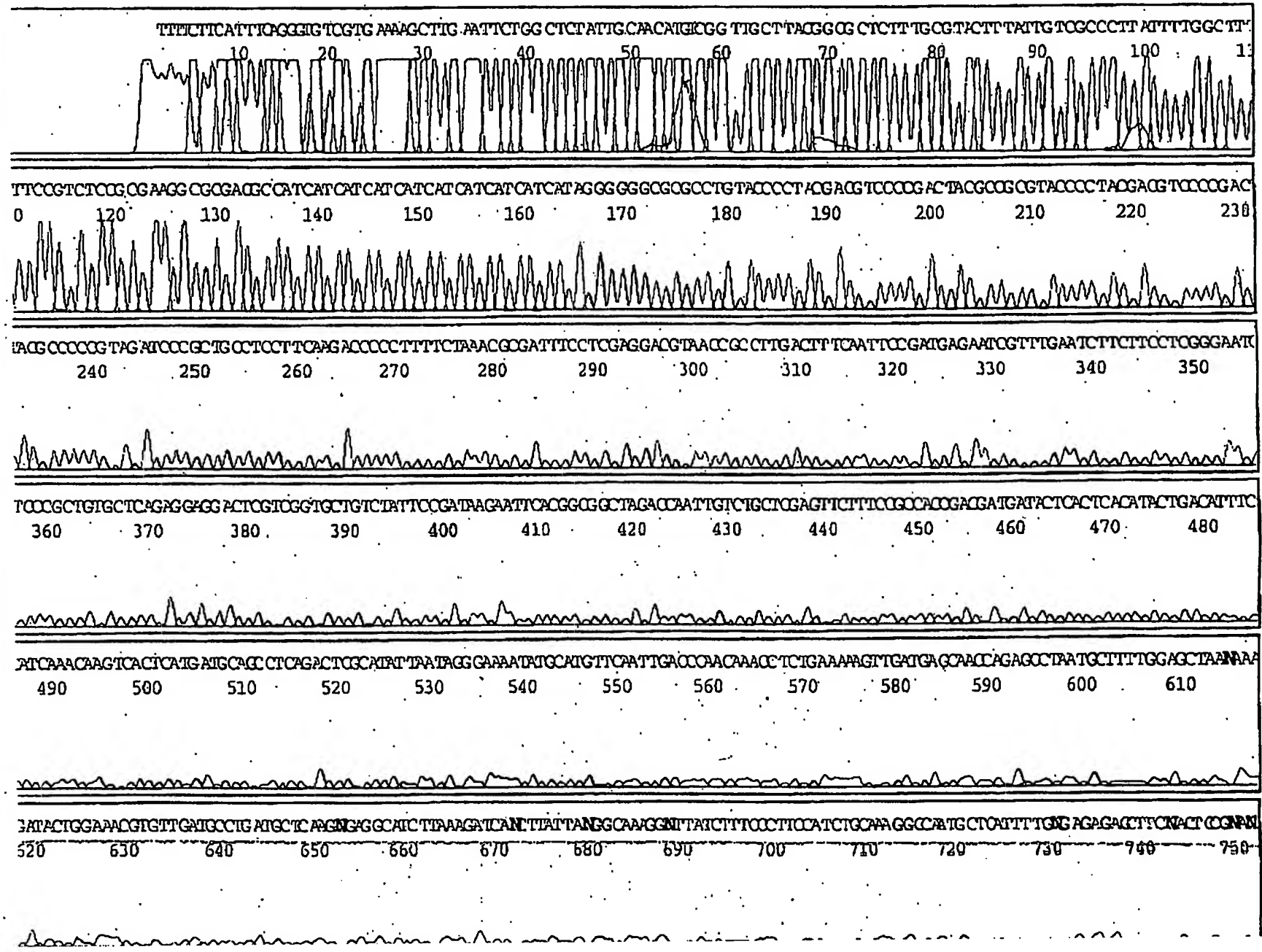
Sample 03
JASON STERLINE
Lane 5

APSg 47780
in pFAC, Forward from SIGNB

Signal G:128 A:144 T:280 C:146
DT3100POP4(BDv3)v1.mob

Points 1566 to 18468 Pk 1 Loc: 1566

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.25[16.25]



ABI PRISM™

Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J533.6P_PERE
BC 1.5.0.0

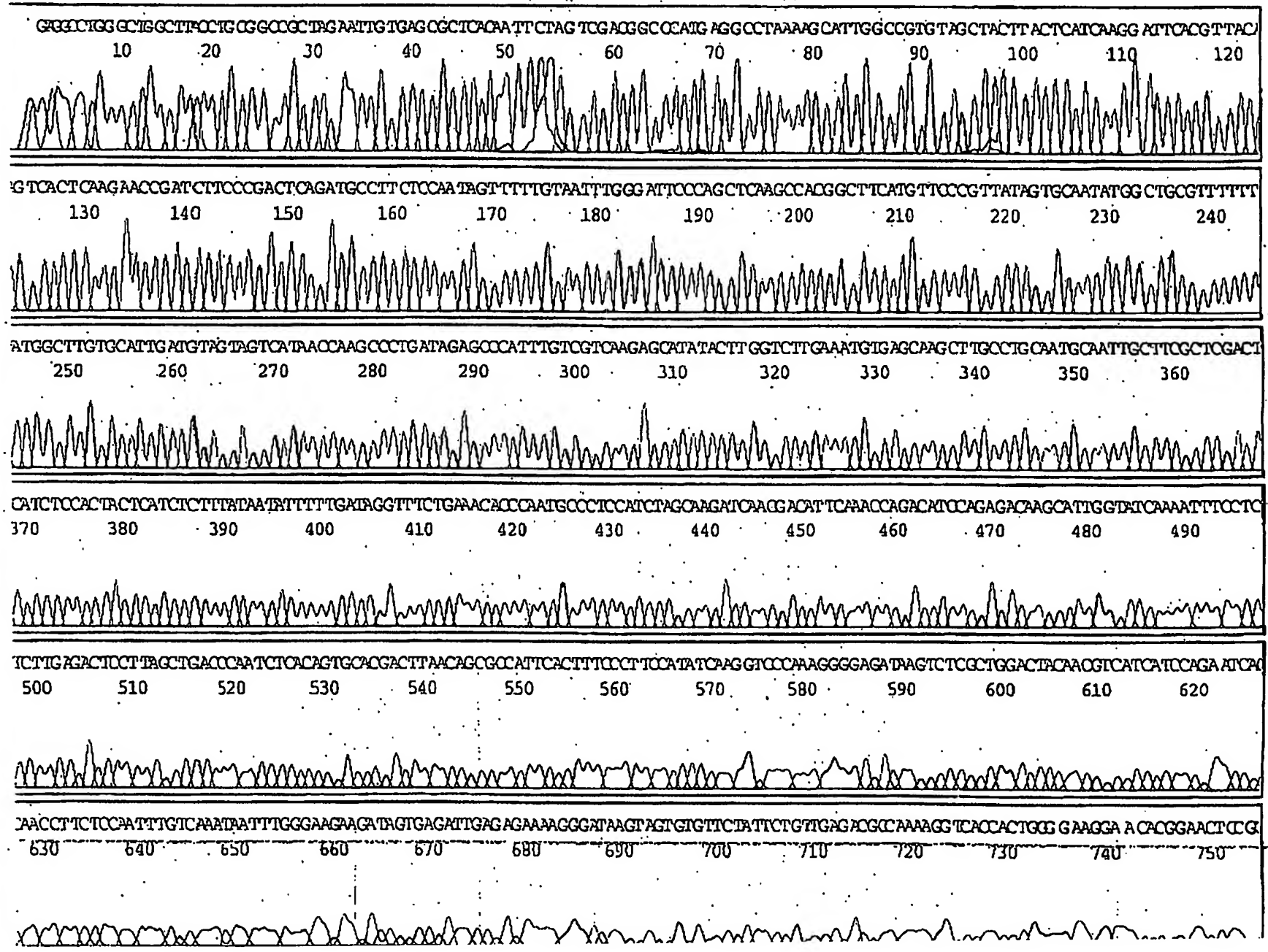
Sample 04
JASON STERLINE
Lane 7

AT2g38650, nP-like
REVERSE, 3' END

Signal G:119 A:149 T:193 C:104
DT3100POP4(BDv3)v1.mob

Points 1810 to 18468 Pk 1 Loc: 1810

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.00(16.00)



ABI PRISM™

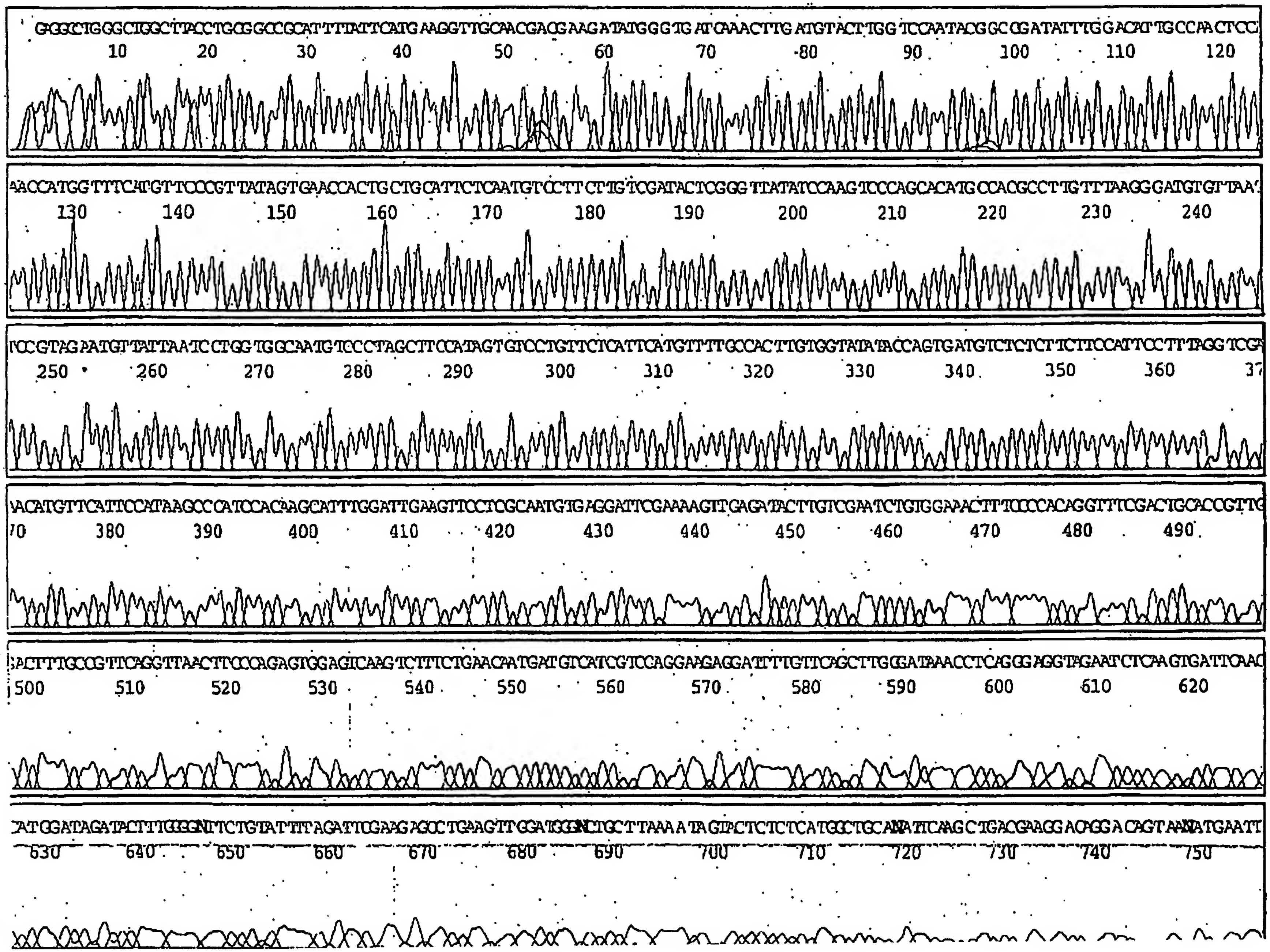
Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J536.7P_PERE
BC 1.5.0.0

Sample 05
JASON STERLINE
Lane 9

At3g61130 in pSAK
REVERSE, FROM 3' END

Signal G:108 A:128 T:190 C:88
DT3100POP4{BDv3}v1.mob
Points 1804 to 18468 Pk 1 Loc: 1804

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.52(16.52)



60445539.02060

ABI PRISM™

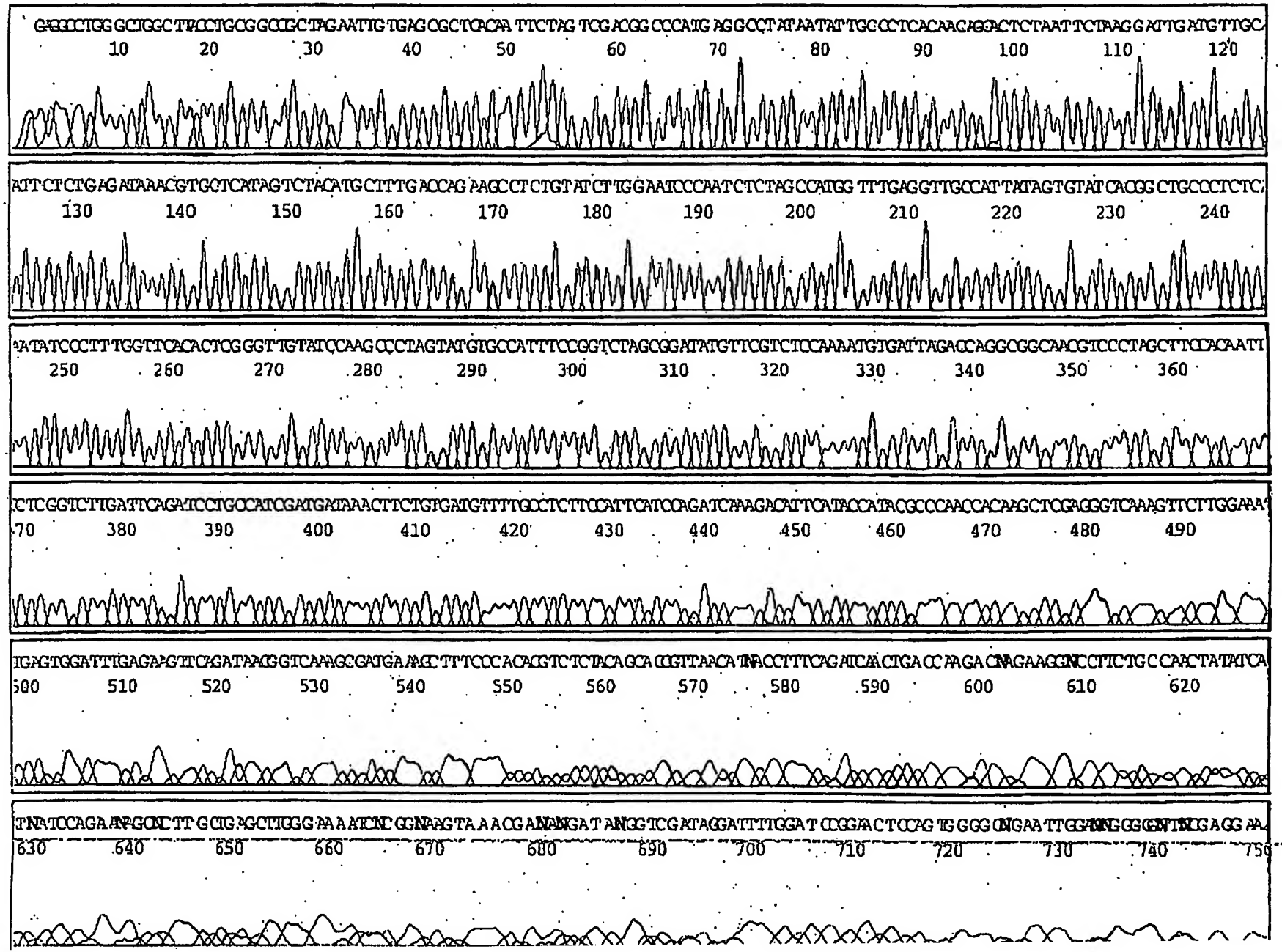
Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J536L.3P_PERE
BC 1.5.0.0

Sample 06
JASON STERLINE
Lane 11

AT5;4778 in PEAK
REVERSE, 3'END

Signal G:104 A:136 T:180 C:94
DT3100POP4(BDv3)v1.mob
Points 1809 to 18468 Pk 1 Loc: 1809

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.25{16.25}



ABI PRISM™

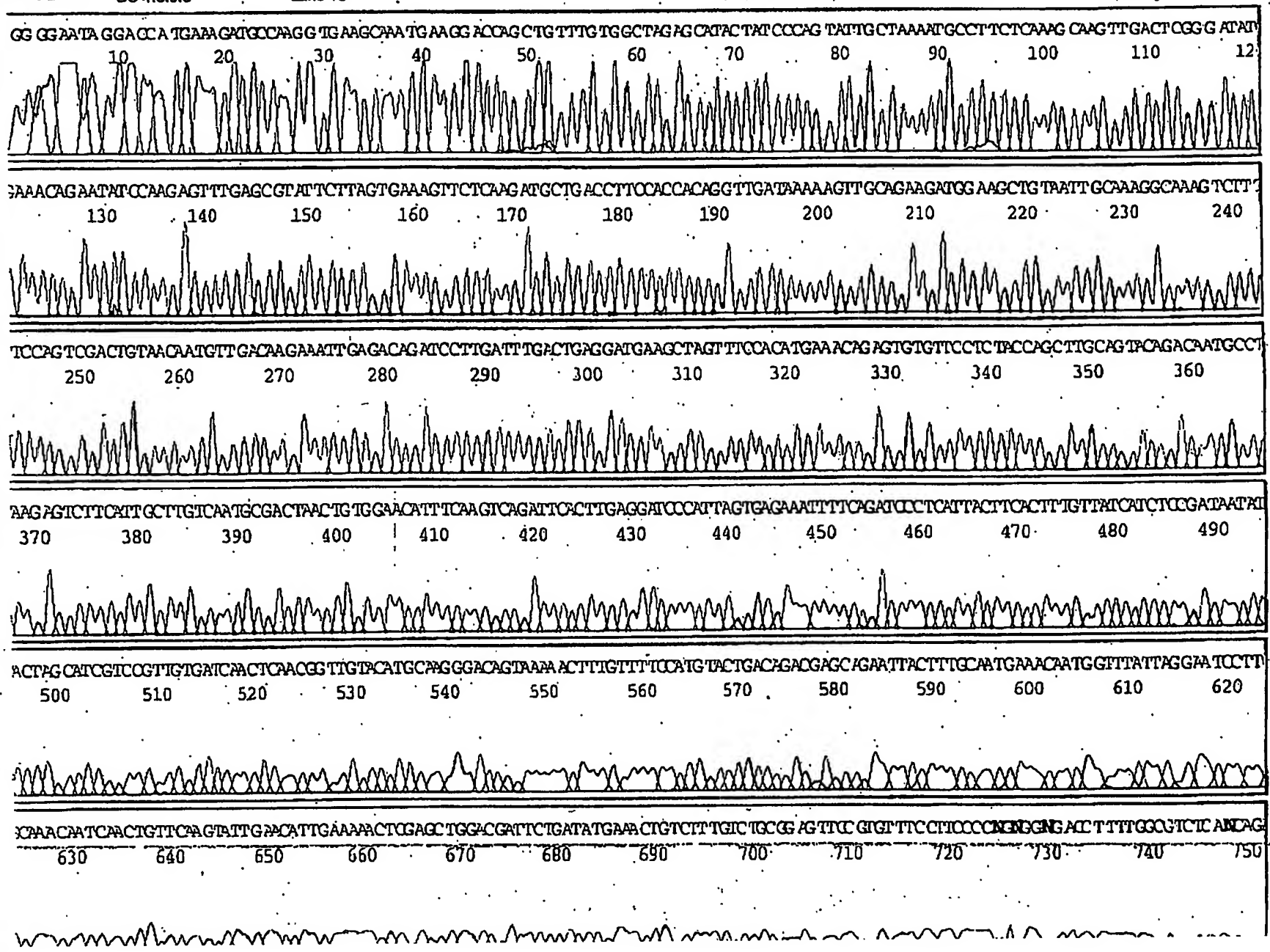
Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J533.6P_331P
BC 1.5.0.0

Sample 07
JASON STERLINE
Lane 13

AT2g38650 - pSAC
FORWARD, INTERIOR PRIMER

Signal G:34 A:65 T:73 C:31
DT3100POP4(BDv3)v1.mob
Points 1798 to 18468 Pk 1 Loc: 1798

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.00(16.00)



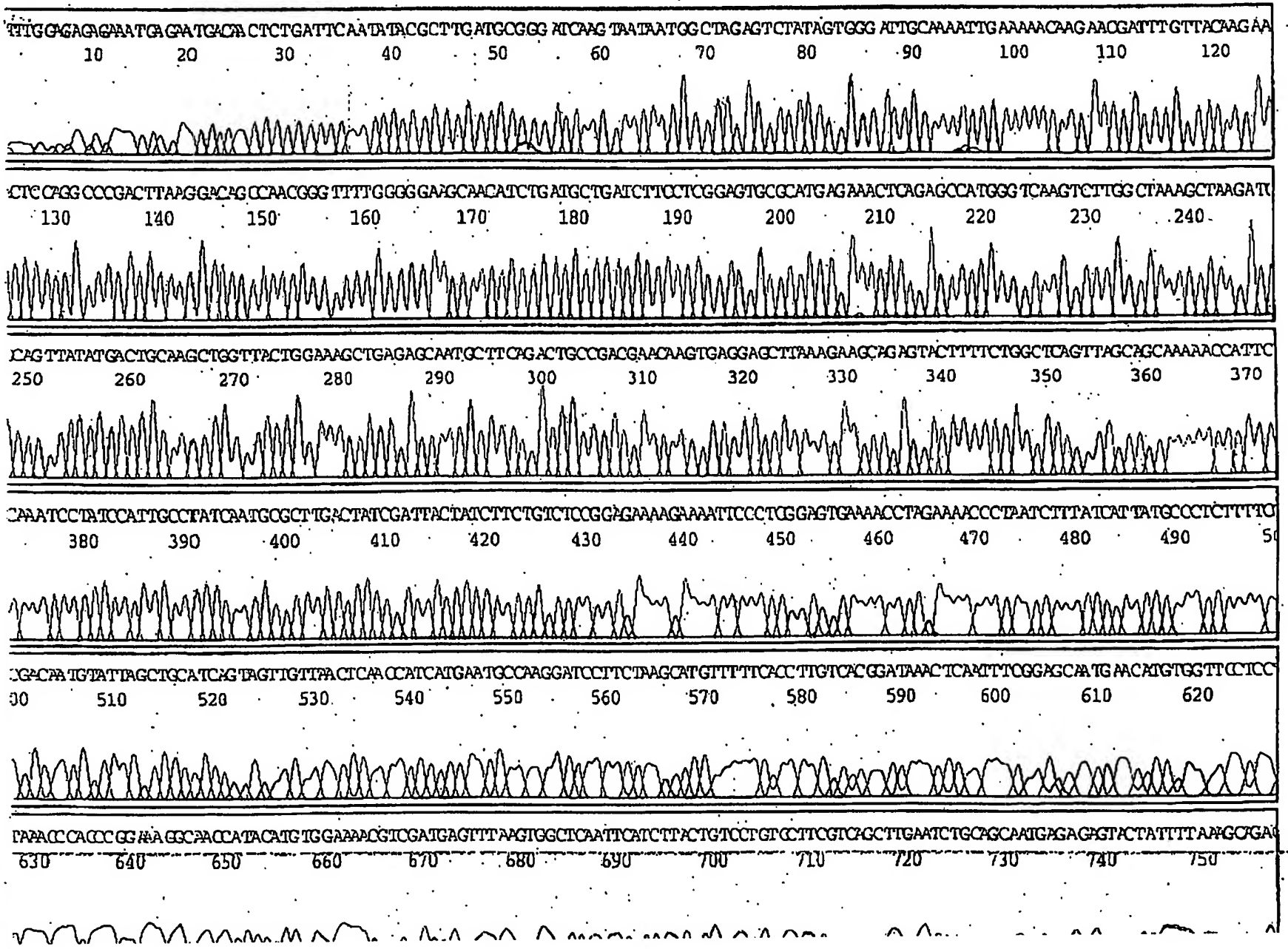
Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J536.7P_361P
BC 1.5.0.0

Sample 08
JASON STERLINE
Lane 15

A#3, 61130 ~ PEAK
FORWARD, INTERIOR PRIMER

Signal G:185 A:282 T:260 C:155
DT3100POP4(BDv3)v1.mob
Points 1824 to 18468 Pk 1 Loc: 1824

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.00(16.00)



Model 3100
Version 3.3
Basecaller-3100POP4_8(N557)J536L_3P_6LIP
BC 1.5.0.0

Sample 09
JASON STERLINE
Lane 2

ABI PRISM

ATS 4778
FORWARD, INTERIOR PRIMER

Signal G:208 A:417 T:395 C:229
DT3100POP4(BDv3)v1.mob

Points 1812 to 18468 Pk 1 Loc: 1812

Tue, Oct 22, 2002 5:16 PM
Tue, Oct 22, 2002 10:55 AM
Spacing: 16.25[16.25]